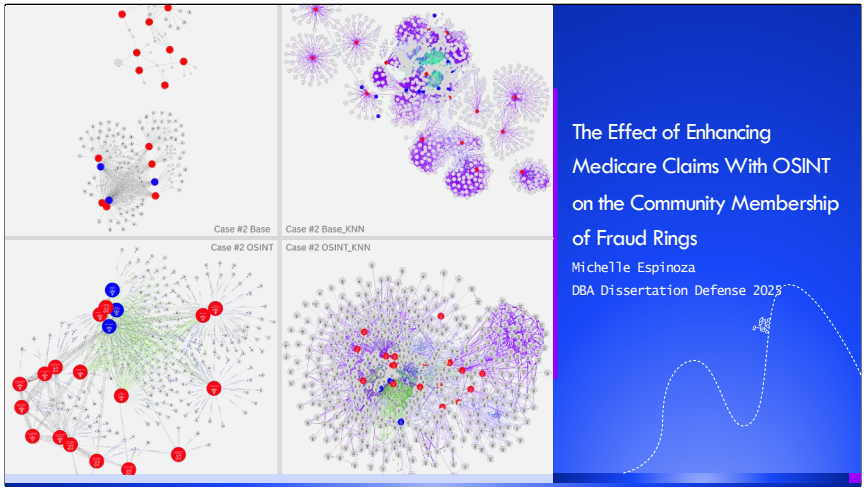
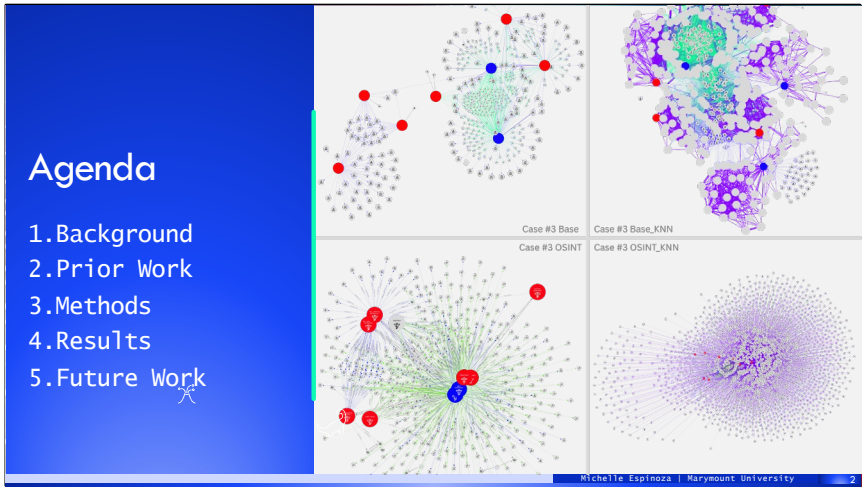


The Effect of Enhancing Medicare Claims With OSINT on the Community Membership of Fraud Rings

Michelle Espinoza

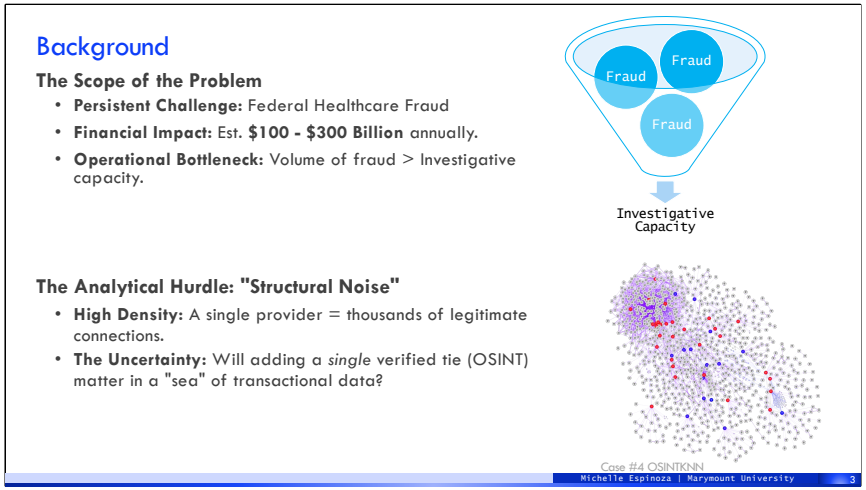
DBA Dissertation Defense 2023



Thank you everyone for being here today and allowing me to present my research.

[click] First I'll give you a background of the study topic and [click] prior work done in this area.

Then I'll move into [click] methods, [click] results, and close out with [click] future work.

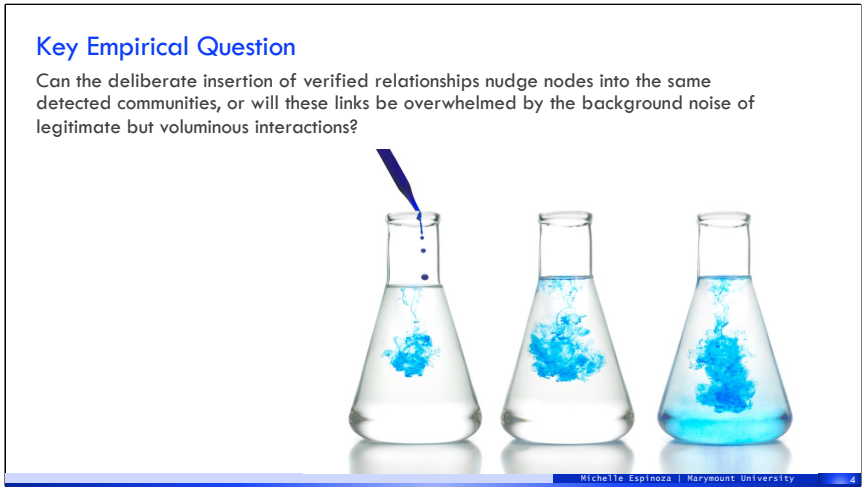


[click] Medicare fraud represents one of the most persistent and costly challenges of federal Healthcare.

[click] Estimates range from one to three hundred billion dollars in annual losses, and investigative agencies continue to face overwhelming caseloads.

[click]This mismatch between the volume of fraudulent activity and the capacity to investigate it highlights the need to explore more efficient analytic approaches.
[click]
At the same time, Medicare claims data are extraordinarily dense.
[click] A single provider may appear connected to thousands of other entities through legitimate billing or geographic relationships.

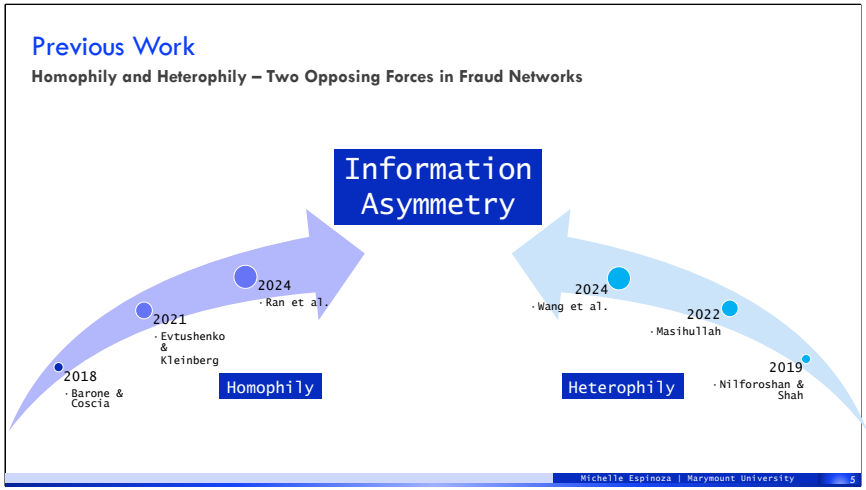
[click] In networks with such high structural noise, it is not immediately clear whether adding a verified relationship between providers will meaningfully influence how community detection algorithms assign clusters.



This uncertainty leads to the central empirical question guiding this research:

[click] Can the deliberate insertion of verified external relationships nudge related providers into the same detected communities, or will these meaningful ties be overwhelmed by the background noise of large-scale claims datasets?

This study addresses that question directly by testing whether targeted OSINT enrichment can improve clustering of known fraud labels.



Prior research presents two opposing forces that shape how fraud networks appear in data: homophily and heterophily.

[click] Homophily is the tendency for similar actors to cluster together.

For example, [click] Barone and Coscia demonstrated homophily within corporate fraud communication networks,
[click] Evtushenko and Kleinberg investigated the Homophily paradox
[click] and Ran and colleagues found homophily in risky prescribing behaviors

These findings suggest that fraudulent actors may naturally cluster in ways that community detection algorithms should be able to identify.

However, other studies show that fraudsters intentionally construct [click] heterophilic relationships to mask collusion. [click]

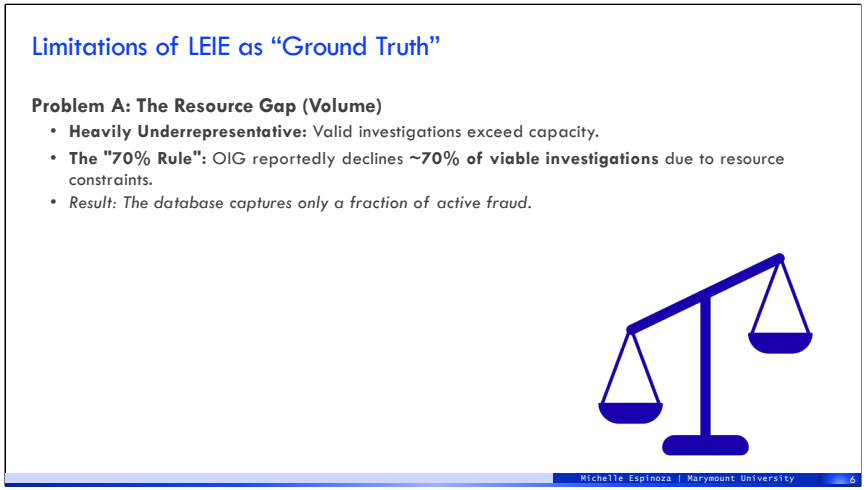
Research by [click] Nilforoshan and Shah, [click] Masihullah, Wang et al., demonstrate that fraudulent networks often embed themselves in dissimilar or benign structures specifically to evade detection.

This creates a paradox: fraud rings rely on collusion—which should make them detectable—yet they simultaneously obscure those ties.

I posit that this tension between homophily and heterophily can be mitigated by reducing [click] information asymmetry.

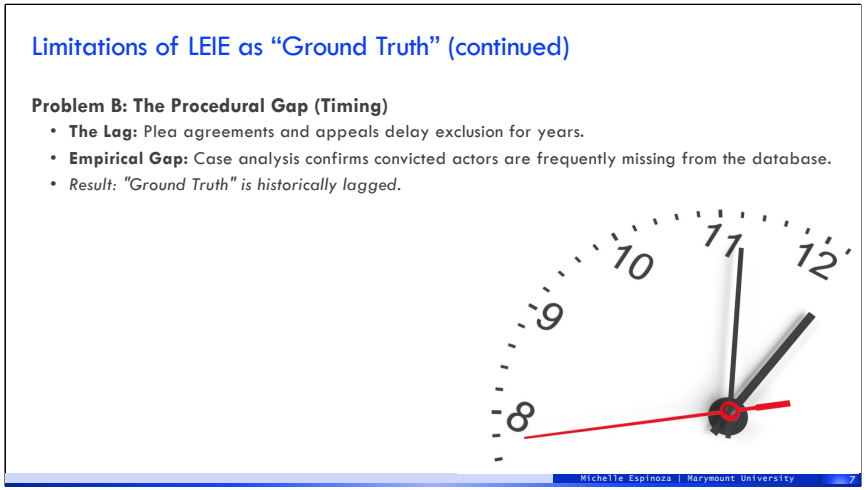
Akerlof's 1970 Information Asymmetry Theory explains how an imbalance of information can lead to market failure.

~~By incorporating OSINT attributes we may be able to reveal the latent homophily that claims-only data fails to surface~~.



A related limitation in the literature is the reliance on the Medicare List of Excluded Individuals and Entities as ground truth for fraud labels.

[click] Although this list is valid it's [click] heavily underrepresentative; thousands of confirmed fraud investigations remain open year over year.
[click] One investigator stated that the Office of the Inspector General turns down 70% of viable investigations each year due to resource constraints.

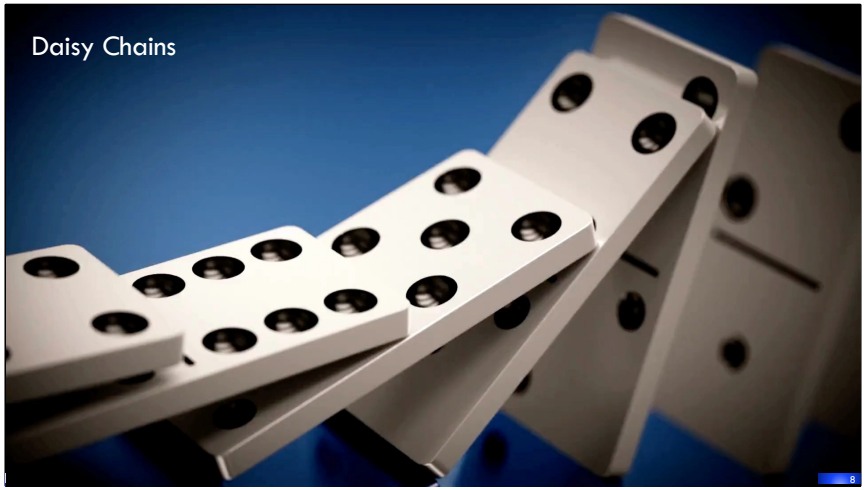


[click]Additionally, many convicted individuals have not yet been excluded due to plea agreements or lengthy legal proceedings which can stretch for several years.

[click] My own case analyses repeatedly showed actors with confirmed fraud histories who were not yet present in the exclusion database.

~~Together, these gaps in the literature and the pervasive investigative challenge highlight the need to explore other analytical approaches.~~

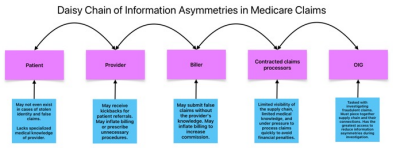
~~This provides the foundation for testing whether OSINT augmentation can reveal hidden structures that traditional claims-only methods fail to capture.~~

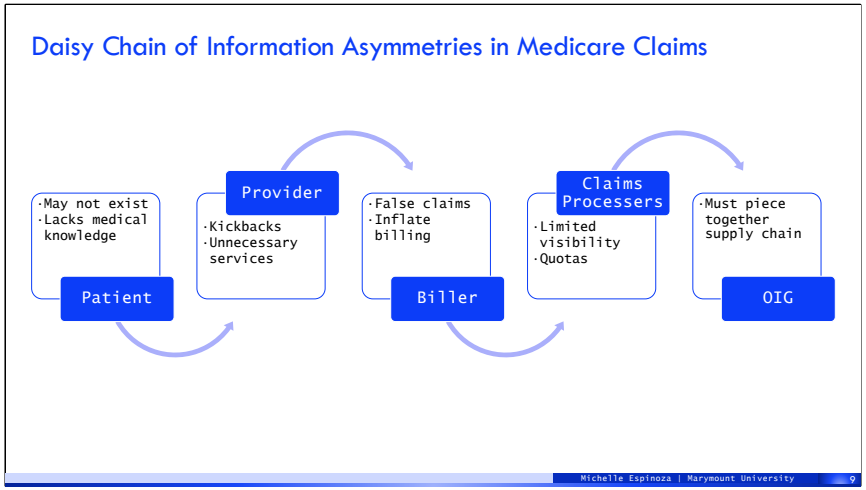


In finance, the term *daisy chain* refers to a coordinated sequence of actors who collude to manipulate market behavior.

In Medicare fraud, a similar daisy chain exists through layers of information asymmetry that create multiple opportunities for moral hazard across the claims lifecycle.

Each actor in the Medicare claims process possesses only a partial view of the system, and fraud can be injected at any stage without full visibility from the others.





[click] Patients may be unaware that their identities are being misused, or may lack the medical knowledge to evaluate the legitimacy of billed services.

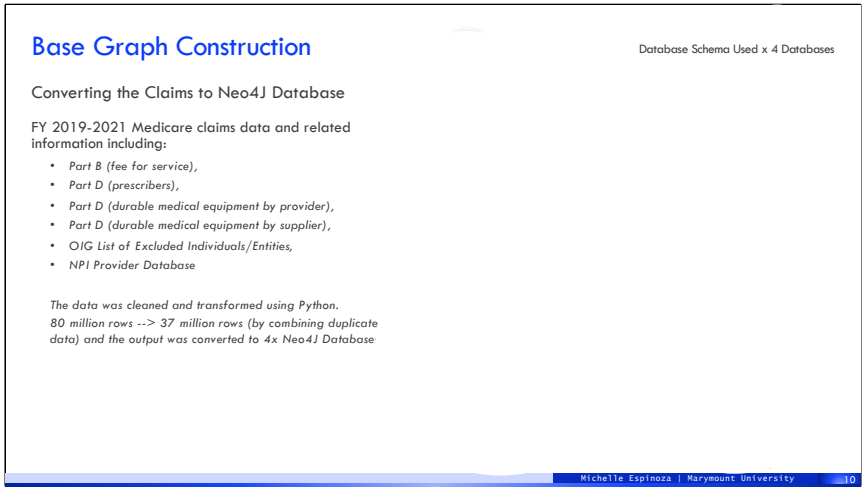
[click] Providers may knowingly or unknowingly participate through referrals, unnecessary procedures, or kickback arrangements.

[click] Billers may submit inflated or entirely false claims with or without the provider's knowledge.

[click] Contracted claims processors, working under strict throughput requirements, typically have limited clinical context and only see fragments of the transaction chain.

And finally, the [click] Office of the Inspector General in their postmortem investigations must reconstruct the entire supply chain after the fact, without timely access to complete information.

The cumulative effect is a cascading loss of visibility, where no single actor holds the complete picture.



The base graph construction for this study uses:
1. Medicare Part B and D claims from fiscal years 2019 through 2021
2. The NPI Provider Registry database and
3. The LEIE database

These datasets represent services, prescriptions, durable medical equipment ordering, provider attributes, and known fraud exclusions.

The raw data consisted of more than eighty million claim-level records.

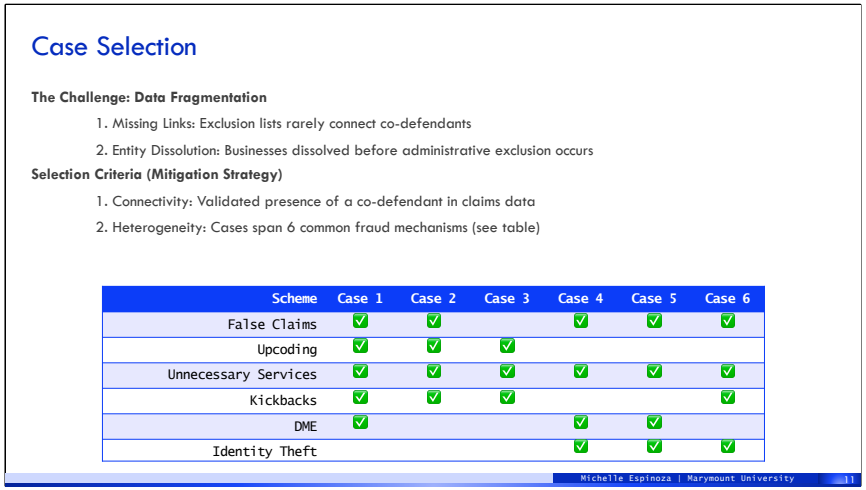
Using Python, these records were cleaned and merged, reducing the dataset to approximately thirty-seven million rows.

The cleaned outputs were then converted into a neo4j database.

On the right side of the slide is the schema used for the graphs.

The nodes represent [click] providers, [click] suppliers, [click] addresses, [click] procedures, [click] medications, [click] equipment and fraud indicators

The edges represent prescriptions, equipment ordering, address and business relationships.



The next step was selecting cases which were representative of common Medicare fraud schemes.

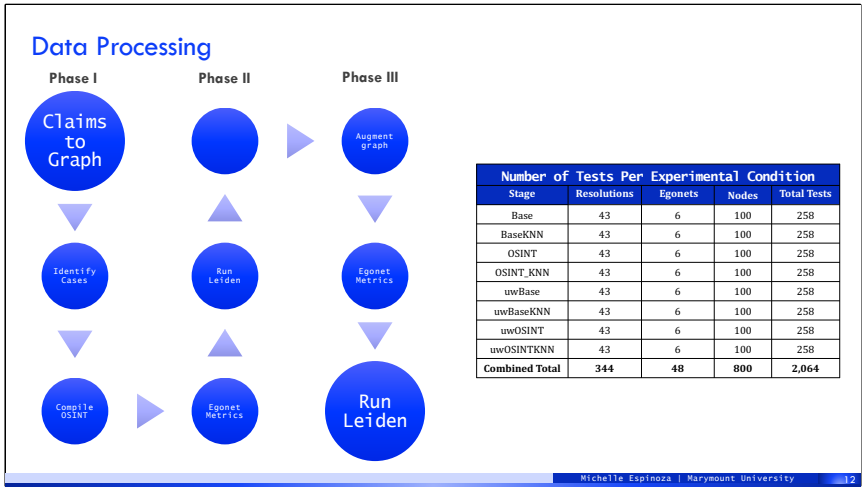
[click] There are inherent limitations in selecting Medicare fraud cases: [click]

1. Exclusion lists don't link codefendants – so even though there may be 10 providers on an indictment- it's possible that 8 of them pled to lesser charges [click]
2. In many instances I found that business entities were dissolved before administrative exclusion occurred

[click] Cases were selected using two criteria: [click]

1. the presence of at least one co-defendant within the claims dataset. [click]
2. Diverse fraud typology [click]
   As shown in the table, the six cases encompass false claims, upcoding, unnecessary services, kickbacks, durable medical equipment fraud, and identity theft
3.
   By selecting cases that vary both in structure and in fraud mechanism, the study is able to meaningfully assess

whether OSINT augmentation improves community detection across a range of heterogenous network structures



This is a high-level summary of the three-phase, eight-step data processing pipeline I followed across all experimental conditions. [click]

After creating the graph databases and [click] identifying cases, [click] OSINT was collected from court documents, corporate records, and medical boards

In phase II, [click] I calculated the network metrics for each egonet before [click] running Leiden community detection at 43 resolutions to establish pre-OSINT community assignments. [click][click]

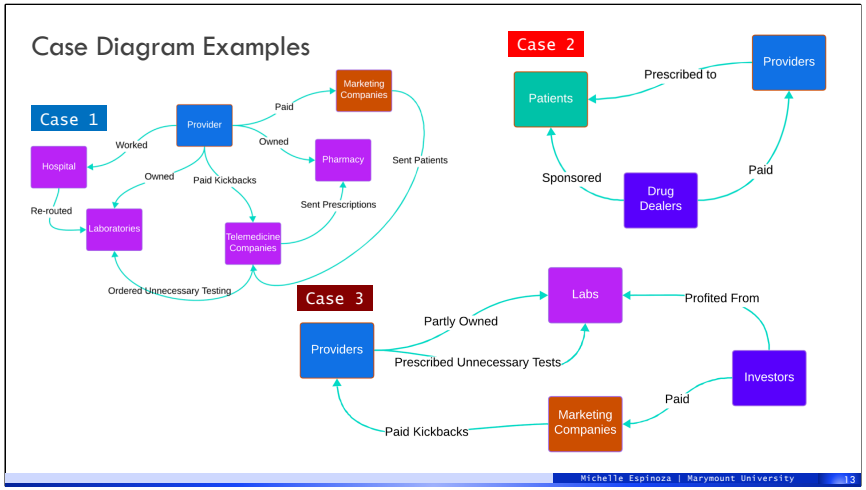
In phase III the graphs were enriched with OSINT and the same metrics and [click] community detection procedures were repeated.

Then I cloned both databases and created FastRP embeddings for Provider nodes

And last, I used Neo4J's K-nearest neighbor's clustering with cosine similarity and recorded the same metrics for each egonet and community membership.

This consistent pipeline ensured that any observed differences in community structure could be attributed to the addition of the OSINT relationships.

The table on the right shows the scale of the experiment: each stage required 43 resolutions across 6 egonets, totaling 258 tests per configuration. Across eight configurations, this produced 2,064 individual Leiden community detection runs.

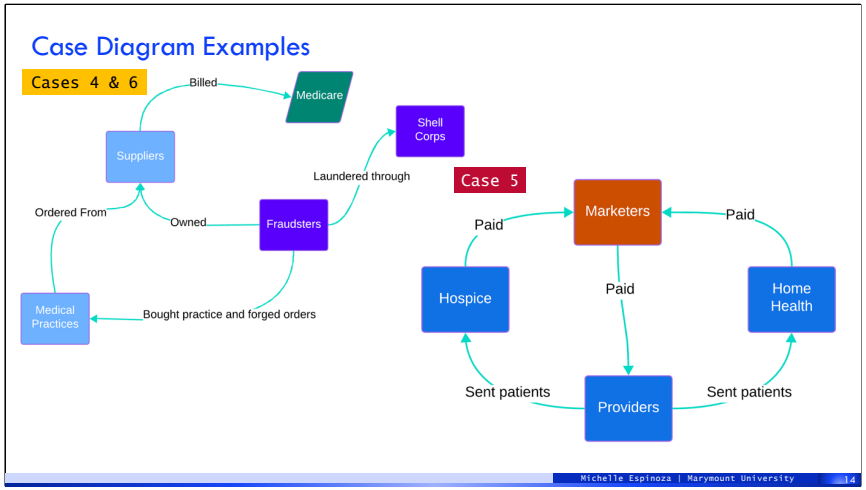


Another thing to note is that each of these cases involves outside entities that would not be surfaced in the graph, even with OSINT (unless external relationships were explicitly added).

These diagrams illustrate the first three case schemes.

1. [click] In case 1, providers paid marketing companies and telemedicine companies to send referrals for unnecessary testing and prescriptions to laboratories and pharmacies owned by the codefendants.
2. [click] In case 2 drug dealers paid providers to prescribe narcotics to patients sponsored by the drug dealers
3. [click] Case 3 is similar to case 1 except in this case, co-investors primarily funneled the patients to complicit providers
4. In these networks, fraudulent behavior typically emerges through the interaction of otherwise legitimate components.

5. And because these relationships are often distributed across corporate structures and geographic regions, they tend not to appear naturally within claims-only data.



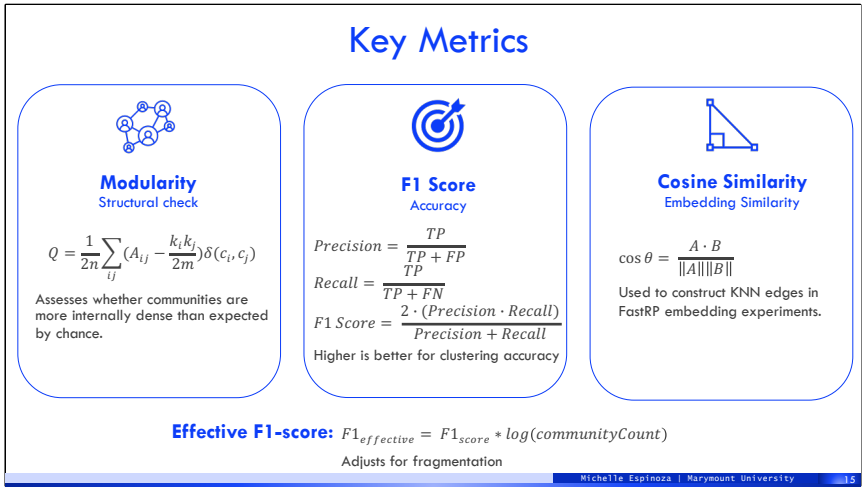
Cases 4 and 6 involved fraudsters purchasing existing medical practices, setting up suppliers and shell companies, then using stolen patient data to fabricate equipment orders that were billed to Medicare.

In a similar example, this summer they indicted a $16 billion dollar fraud ring where all the defendants lived in Russia but were able to purchase existing medical practices and set up companies in the US.

In Case 5, Home Health and Hospice facility owners paid kickbacks to providers, through marketers, to send patients to their facilities denying curative treatment to countless patients in the process.

This demonstrates that no single actor's behavior is sufficient to reveal the broader scheme.

This also reinforces the central premise of the present study: that without external OSINT attributes, claims-only graphs are unlikely to reconstruct the underlying fraud ring, whereas OSINT can supply the relational context needed to improve discovery of these coordinated networks.



Before presenting the results, I want to briefly summarize the key metrics used to evaluate performance in this study.
The first is *modularity*, which serves as a structural quality check.

Modularity measures the extent to which detected communities are more internally dense than would be expected by chance.

It helps ensure that the partitions produced by the Leiden algorithm are meaningful from a network perspective.

The central outcome metric is the *F1-score*, which balances precision and recall.

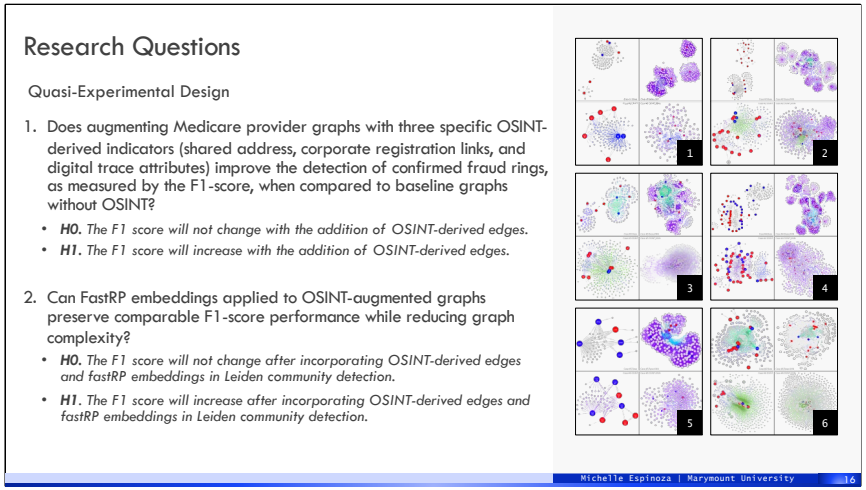
The F1-score therefore provides a single, interpretable measure of clustering accuracy, with higher values indicating better alignment with known fraud ring memberships.

Because large communities can artificially inflate F1 performance, i.e. 100% accuracy across 2 communities is far less meaningful than 100% accuracy across 5,000 communities, I included a custom heuristic the *Effective F1-score*, which rescales the raw F1 by the logarithm of the community count.

Finally, cosine similarity was used for the embedding experiments.

After generating FastRP embeddings for each provider, cosine similarity defined the k-nearest-neighbor edges in the reduced graph, allowing us to evaluate whether simplified structures preserved accurate community detection."

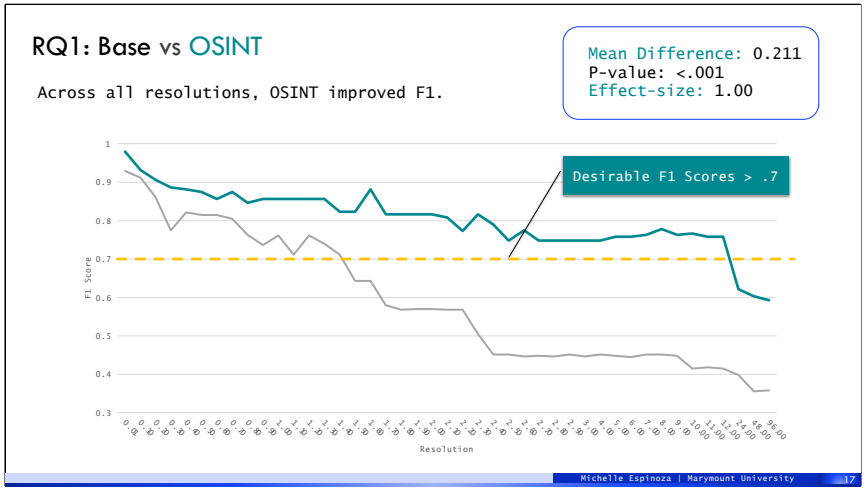
Together, these metrics allow the study to assess both structural integrity and detection accuracy, providing a comprehensive evaluation of how OSINT enrichment affects fraud label community membership



This study uses a quasi-experimental design to address two research questions [click]

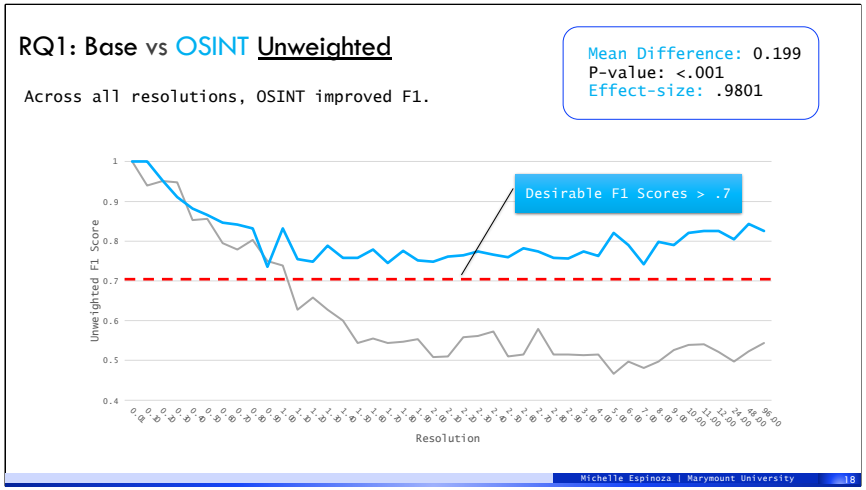
1. Does augmenting Medicare provider graphs with three specific OSINT-derived indicators (shared address, corporate registration links, and digital trace attributes) improve the detection of confirmed fraud rings, as measured by the F1-score, when compared to baseline graphs without OSINT? [click]

2. Can FastRP embeddings applied to OSINT-augmented graphs preserve comparable F1-score performance while reducing graph complexity?



The F1-scores show that OSINT-augmented graphs outperform the baseline at nearly every resolution, with higher accuracy in identifying confirmed fraud ring members. A mean difference of 21%

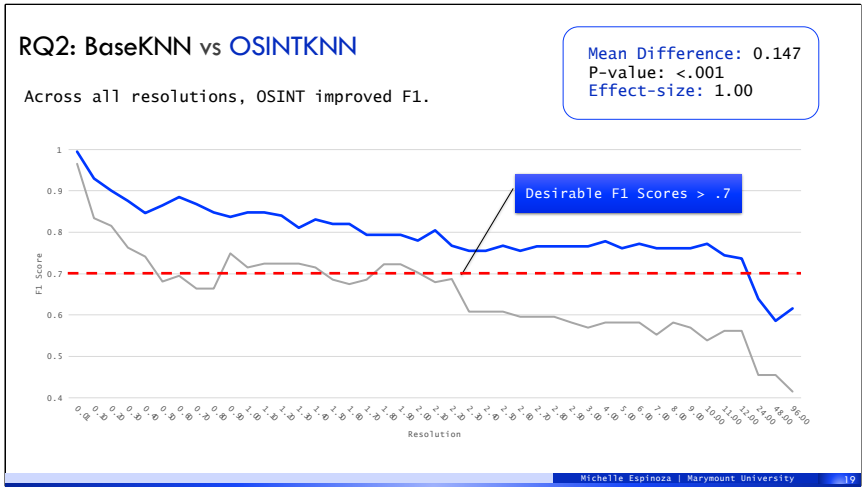
The baseline F1 curve declines sharply as resolution increases, while the OSINT curve remains above the 0.7 threshold for a substantially longer range.



Here I'm showing the unweighted results, which help confirm that the improvements we observe with OSINT are not simply an artifact of the edge-weighting scheme. While both graphs maintained similar clustering accuracy at the lowest resolutions, the base graph's F1 score drops precipitously after a resolution of 1.

The OSINT graph improvement dropped slightly to 19.9%.

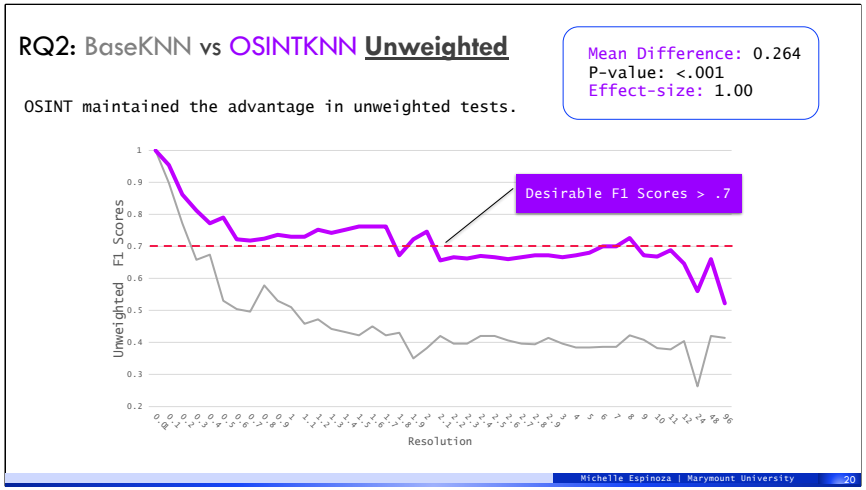
Together, these unweighted models demonstrate that the performance gains from OSINT are robust. Even when all edges contribute equally, the additional OSINT-derived relationships make fraud ring members more likely to cluster together. This confirms that the improvements are driven by the informational content of OSINT—not by the weighting strategy.



This slide shows the results of applying FastRP embeddings to reconstruct similarity-based edges and then comparing F1 performance between baseline KNN graphs and OSINT-augmented KNN graphs.

Across 43 resolution settings, OSINT-KNN produced a mean F1 improvement of approximately **14.7%**. Notably, the performance gap between the base and OSINT-augmented variants narrows in this embedding-based configuration.

One plausible explanation is that FastRP incorporates multi-hop structural information—capturing second- and third-degree relationships—which may reduce the marginal benefit of explicitly inserting OSINT-derived ties. This is consistent with prior observations in fraud networks where actors deliberately obscure direct collusive links, causing much of the relevant signal to appear in higher-order neighborhood structure rather than immediate adjacency.



The unweighted OSINT-KNN graphs outperform the unweighted BaseKNN graphs by a mean of 26.4% across 43 resolutions.

This supports the idea that the original claim-based edge weights strips the base graph of a major source of meaningful signal.

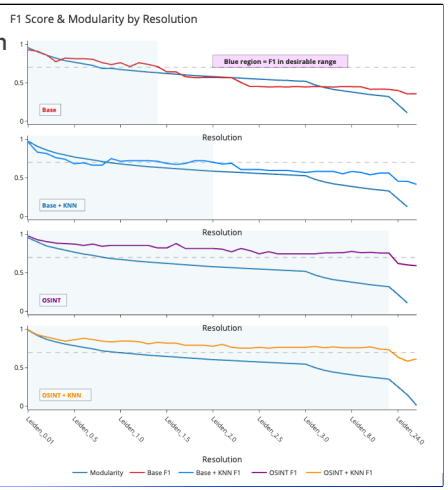
In contrast, the OSINT-augmented graphs still retain meaningful relational structure even without weights. OSINT edges carry categorical, non-transactional information that remains equally strong whether weighted or unweighted. These edges add structurally coherent ties that fraud actors do not expect to be modeled, so they continue to exert influence even after the weights are removed.

F1 Score & Modularity by Resolution

The blue-shaded rectangles show improved F1 scores for OSINT graphs across all resolutions tested.

Mean F1 By Graph

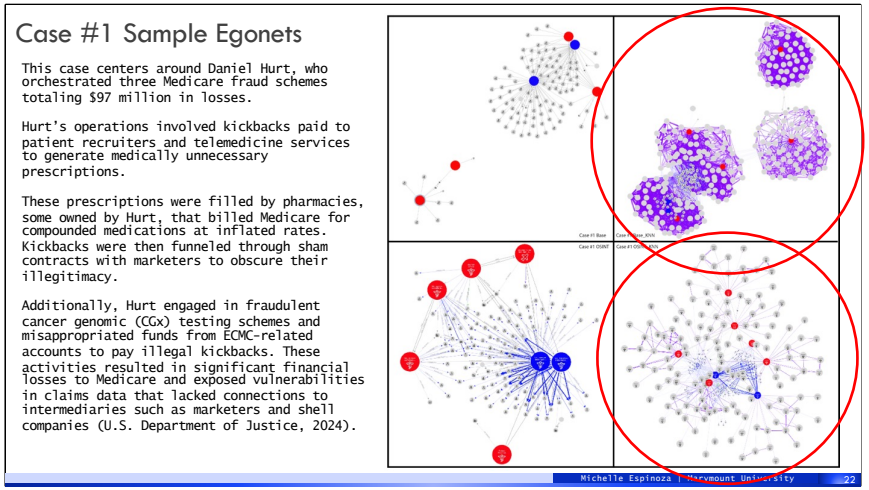
	Weighted	Unweighted	P-Value
Base	.5886	.6204	.36
OSINT	.8012	.8050	.76
BaseKNN	.6496	.4679	--
OSINTKNN	.7953	.7163	--



This slide shows all four F1 score results against modularity.

The shaded blue region highlights the set of resolutions in which the F1-scores are within the desirable range. Both OSINT graphs showed better F1 scores through higher resolutions than the baseline graphs.

This table shows the mean F1 score by graph. Notice that the unweighted Base and OSINT graphs slightly outperformed the weighted graphs, so I ran a t-test and neither difference was significant.



These panels show the Case 1 egonets under each of the four graph conditions.

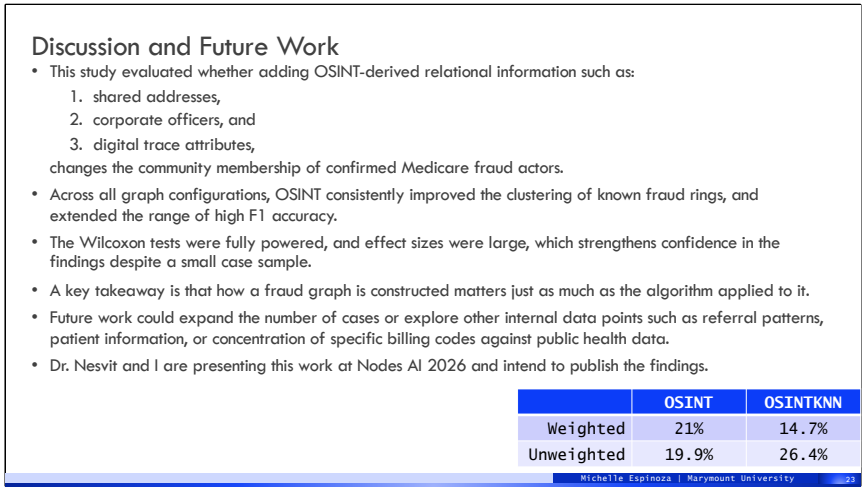
The focal provider orchestrated three simultaneous fraud schemes, resulting in approximately **$97 million in confirmed losses**, and maintained ownership interests in health clinics and compounding pharmacies across multiple states.

One of his schemes included fraudulent cancer genomic testing, which is very prone to fraud, that's likely why the BaseKNN graph in the upper right exhibits tightly packed clusters.

The similarity-based edges pull large numbers of providers into a common neighborhood around high-volume services.

After OSINT augmentation, those inflated clusters break apart. The OSINT edges pull peripheral actors closer to the defendant, and the overall structure becomes more interpretable, with clusters reflecting organizational ties rather

than solely transaction intensity.



[click]This study evaluated whether adding OSINT-derived relational information changes the community membership of confirmed Medicare fraud actors.

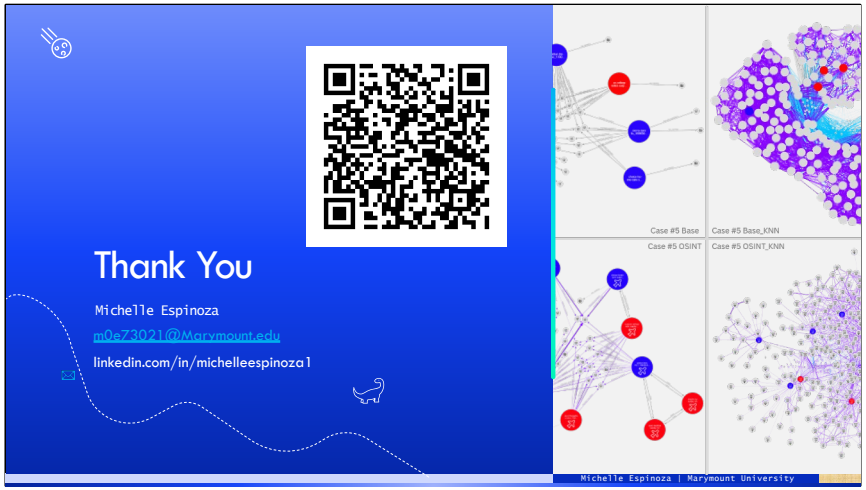
[click] Across all graph configurations, OSINT consistently improved the clustering of known fraud rings, and in many cases extended the range of high F1 accuracy.
[click] The results were robust: the Wilcoxon tests were fully powered, and effect sizes were large, which strengthens confidence in the findings despite a necessarily small sample of labeled nodes.

[click] The results support the idea that fraud is fundamentally a relational and organizational phenomenon, not just an anomaly in billing patterns.

the findings align with complexity-informed views of fraud networks as adaptive, self-organizing systems, thereby situating the study within a systems-level paradigm of fraud detection.

[click] Future work could expand the number of cases or explore other internal data points such as referral patterns, patient information, or concentration of specific billing codes against public health data.

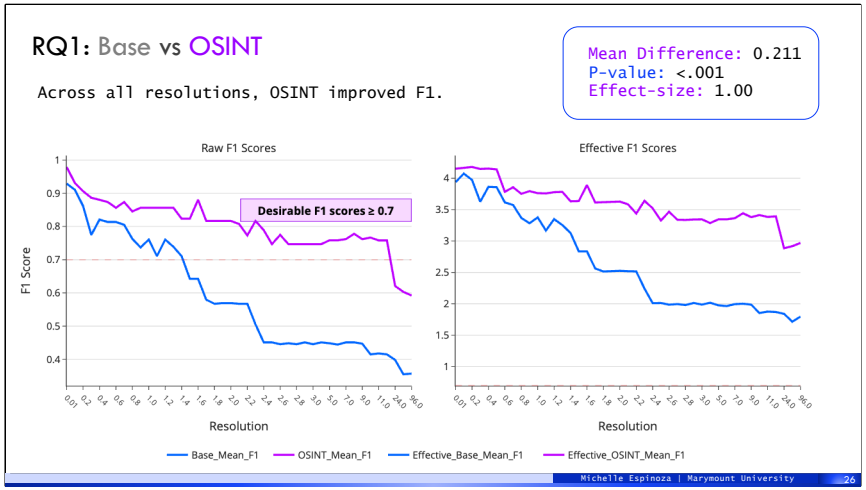
[click] Dr. Nesvit and I are presenting this work at Nodes AI 2026 and intend to publish the findings.



That concludes my presentation. Thank you again for allowing me to present my research today, and a special thank you to Dr. Nesvit for seeing the light in my work when no one else did, and for your patience, guidance, and expertise in helping me refine it into a focused and digestible form.

And on that note – I think we can move onto questions.

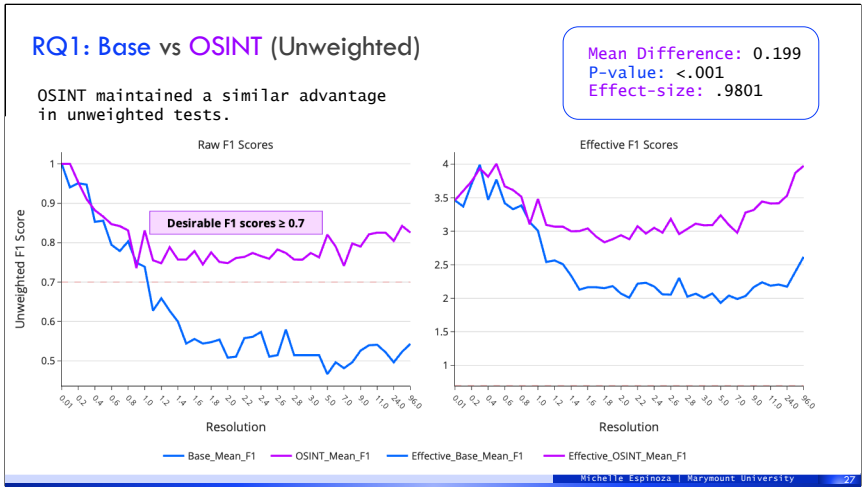
Appendix



On the left, the raw F1-scores show that OSINT-augmented graphs outperform the baseline at nearly every resolution, with higher accuracy in identifying confirmed fraud ring members. A mean difference of 21%

The baseline F1 curve declines sharply as resolution increases, while the OSINT curve remains above the 0.7 threshold for a substantially longer range.
On the right, the heuristic Effective F1-scores show the same pattern after adjusting for fragmentation.

OSINT-augmented graphs maintain higher effective F1 values across all resolutions, indicating that the improvement is not an artifact of community size.

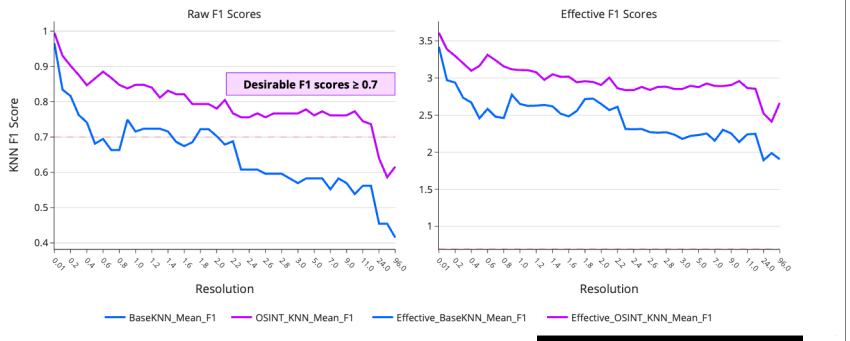


Here I'm showing the unweighted results, which help confirm that the improvements we observe with OSINT are not simply an artifact of the edge-weighting scheme. While both graphs maintained similar clustering accuracy at the lowest resolutions, the base graph's F1 score drops precipitously after a resolution of 1.

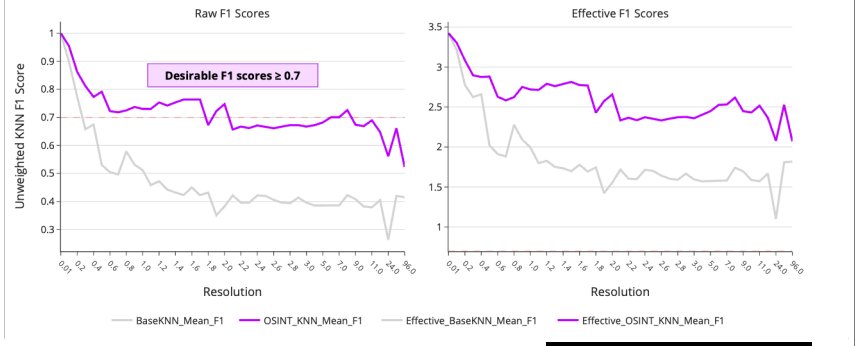
The OSINT graph improvement dropped slightly to 19.9%.

Together, these unweighted models demonstrate that the performance gains from OSINT are robust. Even when all edges contribute equally, the additional OSINT-derived relationships make fraud ring members more likely to cluster together. This confirms that the improvements are driven by the informational content of OSINT—not by the weighting strategy.

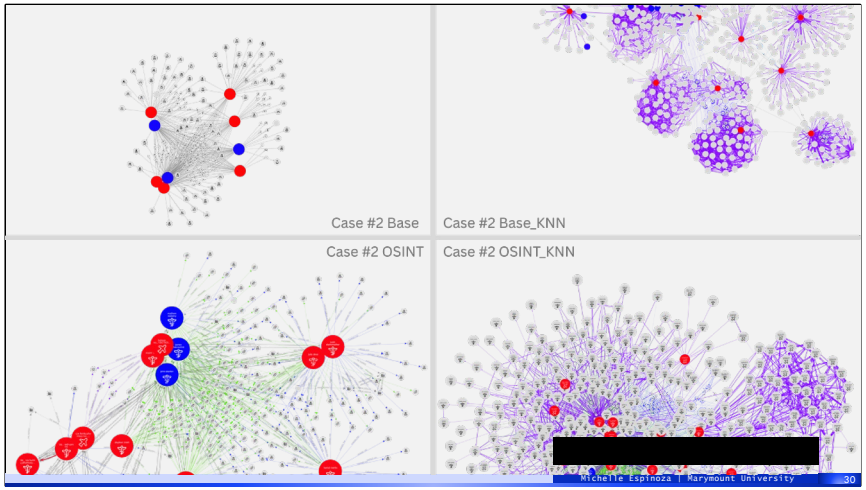
BaseKNN vs OSINTKNN



Unweighted Base KNN vs OSINTKNN



[redacted]


Case #2 Base
Case #2 Base_KNN
Case #2 OSINT
Case #2 OSINT_KNN
[redacted]
Michelle Espinoza | Marymount University
30