Data analytics for Cyber security

-Cybersecurity through Time Series and Spatial Data

-

Vandana P. Janeja

Outline

Spatial and Temporal Data

- Spatial Data
- Spatial Autocorrelation, Spatial Heterogeneity
- Temporal Data
- Temporal Autocorrelation
- Temporal Heterogeneity

Some Key Methods for Anomaly Detection in Spatial and Temporal Data

- Spatial Anomaly Detection
- Spatial Neighborhood
- Temporal Anomaly Detection
- Temporal Neighborhoods

Cybersecurity through Spatiotemporal Analysis

- Spatial Anomalies in Cybersecurity
- Neighborhood-Based Anomalies
- Spatiotemporally Clustered Anomalies
- Insights through Geo-Origination of IP Addresses

Temporal Behaviors in Evolving Networks

Spatial and Temporal Context

- Spatial data refers to geo spatial data in the context of cybersecurity
- The spatial context comes from the origination of the attacks or sources of attacks
- The network layout can also have a spatial context
- One example of spatial context in cybersecurity is the geo-origination of anomalies such as massive spam originating from certain spam hubs around the globe
- Temporal context refers to the time reference of an attack or network traffic in general over a period of time
- Temporal context refers to how network traffic evolves over time or how certain trends over time could indicate threats

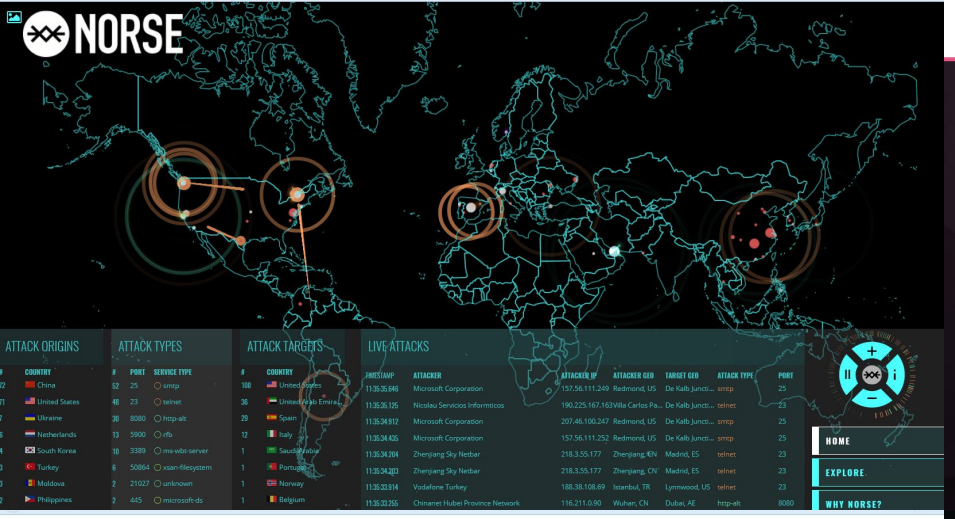
Spatial and Temporal Data

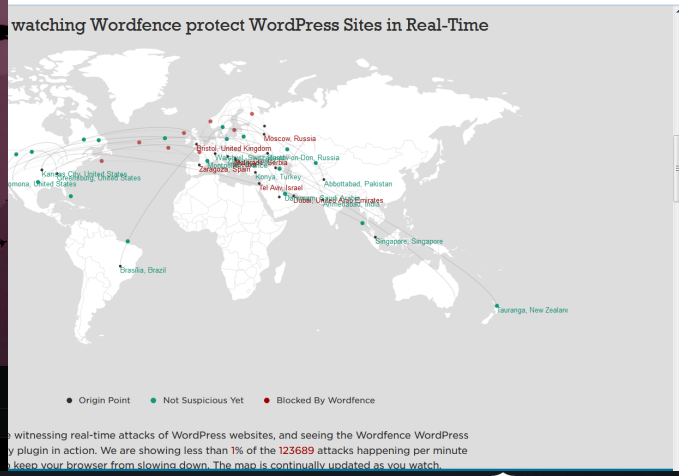
Spatial Data

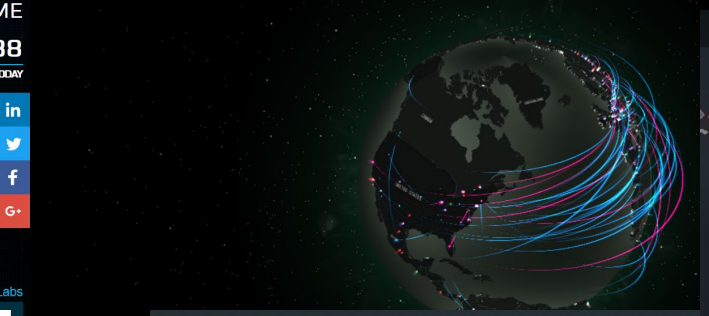
- Spatial data consists of data pertaining to space, multidimensional points, lines, rectangles, polygons, cubes and other geometric objects
- Spatial data has large volumes so much so that it may require an infinitely large database to capture the real world precisely
- Spatial data not only poses challenges due to the type of data and volume but also due to the correlation between the objects and their spatial neighbors with a neighborhood
- There can be relationships between spatial objects like topological relationships (e.g. disjoint, meet, overlap, etc.), direction relationships (e.g. North, South, above, below, etc.), metric relationships (e.g. distance greater than) or complex relationships (e.g. overlap with a certain distance threshold in the north), which are based on these basic relationships. Moreover, the data can depict temporal changes, which is inherent in the data itself
- For the purposes of cybersecurity we focus on geo referenced data, which includes spatial location in terms of latitude, longitude and other information about spatial and non-spatial attributes
- This essentially means that there is some preprocessing on the raw data to extract the geo referenced attributes and some other non-spatial attributes such as geo-political attributes, other regional attributes affecting the network trends such as network bandwidths etc.

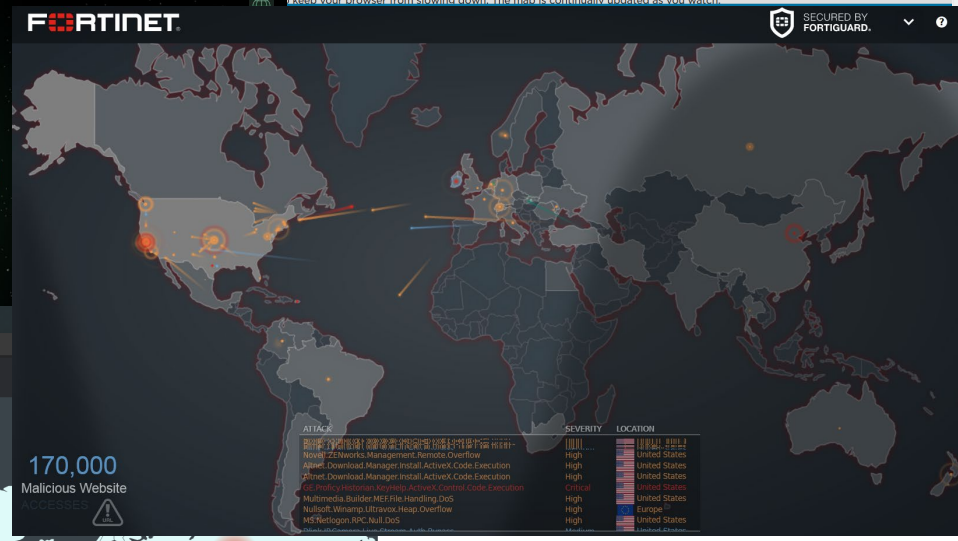
Cyber Attack Maps Which Capture Information In A Geo Referenced Manner

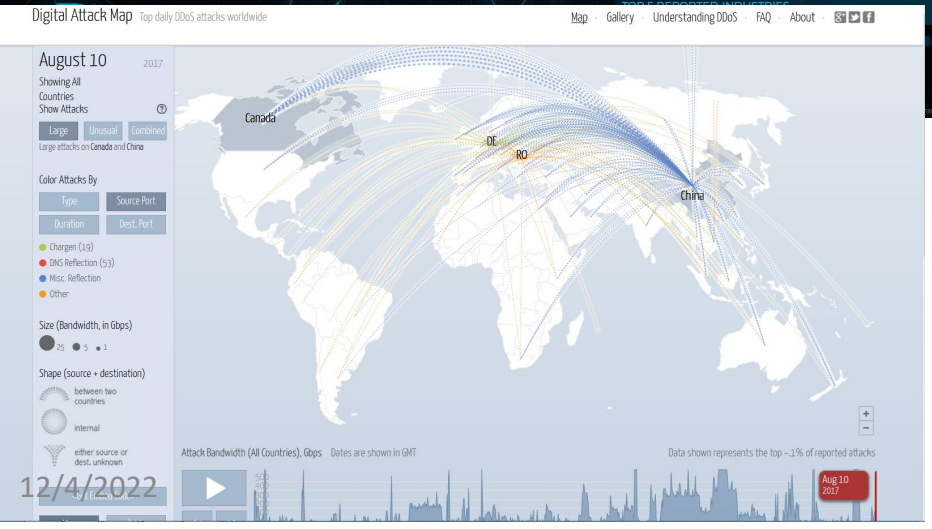
Cyberattack map	Link	Cyber attack information
Norse	http://map.norsecorp.com/#/	Attack type, attack origin and target country, live attack map
Checkpoint- threatmap	https://threatmap.checkpoint.com/ThreatPortal/livemap.html	Attacker, targets, top countries attacked and top attackers
Fireeye	https://www.fireeye.com/cyber-map/threat-map.html	Source, destination, top industries targeted
Kaspersky	https://cybermap.kaspersky.com/	Interactive based on data sources, threat types
Digital attack map	http://www.digitalattackmap.com/#anim=1&color=0&country=ALL&list=0&time=17388&view=map	Source destination, duration, size
Akamai	https://www.akamai.com/us/en/solutions/intelligent-platform/visualizing-akamai/real-time-web-monitor.jsp	Network attacks by region
Wordfence	https://www.wordfence.com/	Wordpress plugin, blocking attacks

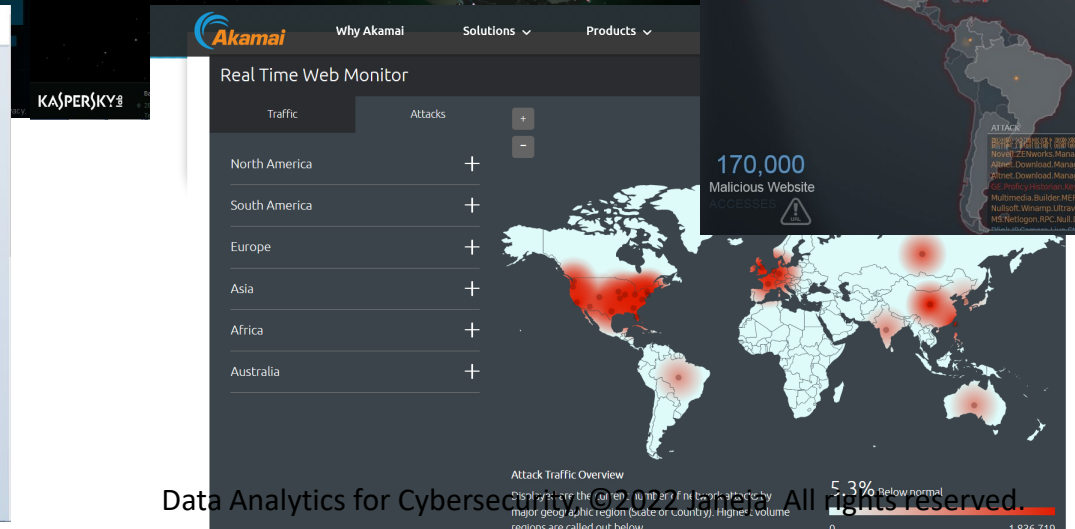










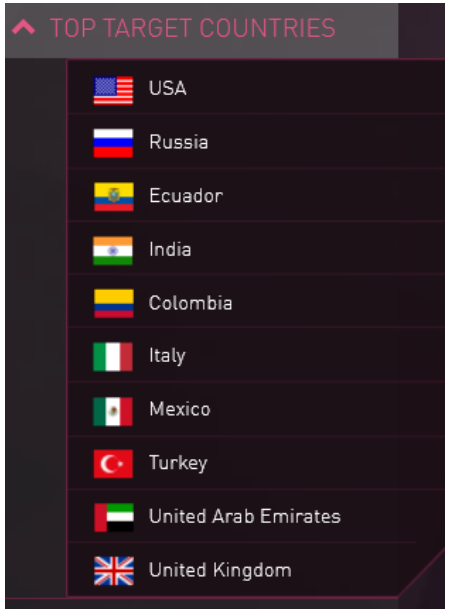


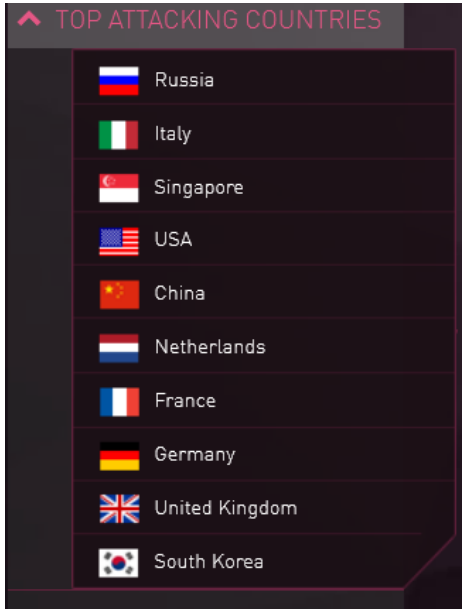
Spatial Autocorrelation

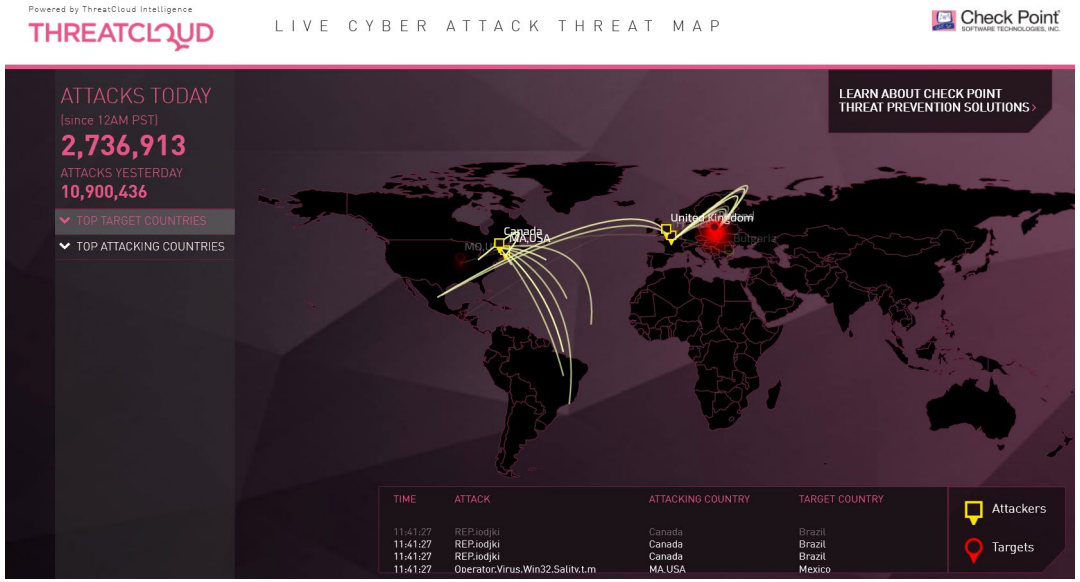
- Autocorrelation refers to the correlation of a variable to itself
- Spatial autocorrelation refers to the correlation of the variable with itself in space. It can either be positive (spatial clusters for high-high or low-low values) or negative (high-low or low-high values)
- Everything is related to everything else but nearby objects are more related than distant objects (Tobler 1970)- core of spatial autocorrelation
- Any statistical measure (such as Moran's me, Geary's ratio or Joint Count of spatial autocorrelation relies upon this (Tobler's) first law of geography

Spatial Heterogeneity

- Spatial Heterogeneity is the spatially varying autocorrelation, where autocorrelation can be positive or negative
- It refers to the variation in the region that may generate characteristic spatial patterns, which could be due to the underlying geographical processes and geographical features in the region
- This behavior can be understood in different scales of spatial variation, including large (macro variation as autocorrelation) and medium/small (meso variation as heterogeneity), as well as error (independent noise or anomalies.
- Based on the concept of spatial autocorrelation and heterogeneity, it may be possible that the geo spatial trends within United States may be similar due to spatial autocorrelation, however, there may also be inherent variability due to heterogeneity
- Let us consider the number of attacks coming into the United States
- It is clear that every location in the United States will not be targeted with the same frequency
- So while the number of incoming attacks are generally high for the United States, depicting spatial autocorrelation, there is also a spatial heterogeneity due to the relative variation of the attacks to different locations inside the United States
- The example shows that the United States is the top targeted country however, the number of attacks on the east coast are much larger than the Midwest due to the richness of targets in the east coast
- Even though there is a bigger class of locations, for a more accurate analysis the locations will need to be subdivided to better evaluate and better allocate resources







Temporal Data

- Let us consider a location or an entity, for example, identified by an IP address
- This location or entity may have measurements for a variable separated by a regular interval of time
- For example, the traffic information originating from the IP address is generated continuously over time (separated by seconds or milliseconds)
- This traffic is linked to the geolocation generating the traffic
- This data by nature is temporal as it is time ordered and each measurement is separated by a time tick (seconds, milliseconds or Nanoseconds)
- The attribute or variable being measured may be the data gram length (dgmLen) or size of the packet
- There is also a possibility that the object or entity is the network itself and the variable being measured over time is the number of incoming or outgoing connections to see the health of the network in terms of the network traffic
- Similar to spatial data, temporal data also has issues of autocorrelation and heterogeneity.

Temporal Autocorrelation

- Two instances of an object in time occurring one after the other (reflected as two attribute values for this object) have a temporal relationship of adjacency
- Since these instances occurred one after the other, they are possibly similar or are related in terms of their values (quantified as Temporal Autocorrelation)
- The time difference between the values of the attribute is referred to as lag
- For example, for the variable ‘packet size’ for an IP address the lag is one millisecond. An autocorrelation coefficient (such as Pearson’s ’r’) measures the correlations between temporal values a certain temporal distance apart
- Here autocorrelation measures the correlation (positive or negative) of the data variable to itself at a given time lag

Temporal Heterogeneity

- The overall autocorrelation of a temporal data sequence may be positive and high
- However, there may be an underlying processes that may bring about some patterns, which are averaged out over larger time periods
- Although temporal mining accounts for issues such as seasonality it does not account for the bigger challenge of heterogeneity
- Heterogeneity essentially is the property of the temporal data due to which it behaves differently even in close time proximity A temporal sequence, which depicts an ascending trend may give a high value for autocorrelation when considered for larger time window, on close observations we may find pockets of variations in this data due to the inherent variability of the data
- If autocorrelation exists in temporal data it can be assumed that data is coming from the same underlying process for the points in a time window
- Heterogeneity indicates that even though the data may be autocorrelated it may be originating from multiple processes depicting micro scale variations

Some Key Methods for Anomaly Detection in Spatial and Temporal Data

Spatial Anomaly Detection

- Anomaly detection deals with discovering non-trivial and intriguing knowledge, in the form of unusual patterns, objects and relationships
- Such a discovery works on the principle of identifying anomalies with respect to the similarly behaving partitions in the data
- Accounting for heterogeneity within the data in creating these partitions is critical to accurately identify anomalies
- To understand the need for such a refined anomaly detection the current approaches to anomaly detection in spatial and temporal data are discussed with an emphasis on identifying the correct partitioning, referred to as neighborhood discovery, in spatial and temporal data
- The neighborhoods can also be used as a characterization such that objects which deviate substantially from this characterization can be identified as suspicious
- This process is similar to anomaly detection using clustering, the key difference is that the preprocessing to form the similar objects into a neighborhood is different for spatial and temporal data due to the properties of the data outlined above
- The anomaly detection process is similar for all data types the key difference is the pre-processing based on different types of data

Spatial anomaly detection

Spatial outlier detection: individual object or a set of objects can be outliers with respect to neighboring objects

Anomalous window detection: a set of contiguous spatial objects are outliers with respect to the entire data

Spatial Anomaly Detection

- A spatial outlier is an object which is deviant in its behavior from its neighboring objects
- This behavior is quantified in terms of significant difference in attribute values of the object with similarly behaving objects in a certain proximity (neighbors based on spatial relationships)
- Several techniques consider neighborhoods that are based primarily on spatial relationships, thus, they only account for autocorrelation, but ignore spatial heterogeneity
- The neighborhood formation is not order invariant-a neighborhood based on cardinality (number of objects in the neighborhood), will lead to different neighborhood formation and outlier detection outcomes with different starting points
- The outlier detection considers the deviation in terms of a single attribute only or Euclidean distance between multiple attributes and not interrelationships between attributes

Spatial Anomaly Detection

- Anomalous Window Detection deals with the discovery of contiguous set of objects forming an unusual window
- In most clustering techniques (such as DBSCAN, OPTICS, the focus is on the identification of the major groupings in the data, as opposed to unusual ordered groupings
- Similar to clustering, outlier detection identifies individual outliers and does not quantify the degree of unusualness of an outlier
- Scan statistic approaches detect a group of objects behaving anomalously with respect to all the objects in the dataset, and detect quantifiable anomalous behavior in multiple attributes in relation to one another
- Scan statistic techniques (tempora, spatio-temporal) identify clusters that are unusual with respect to the rest of the data

Spatial Neighborhood

•The goal of this step is to identify spatial neighborhoods accommodating both spatial autocorrelation and heterogeneity in the region

•The neighborhood should not be identified simply by the change in geospatial features or solely on the basis of spatial proximity determined using spatial relationships such as adjacency of objects

•Generate the immediate neighborhood of an object, called as **micro neighborhood**

• This captures the entire knowledge about the immediate neighborhood of this object such as the presence or absence of spatial features (e.g.: proximity to geo political events or prior known cyber attacks from this location) in its proximity and attributes of the object

•This will allow to associate features and attributes to these objects which can be used to determine similarity or dissimilarity across spatial objects

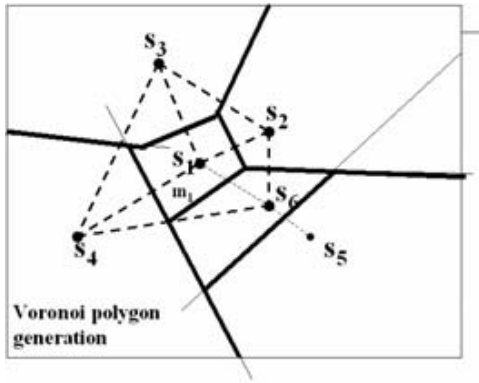
• Identify **spatial relationships** (such as adjacency) using the spatial attributes of the micro neighborhoods, to accommodate for autocorrelation

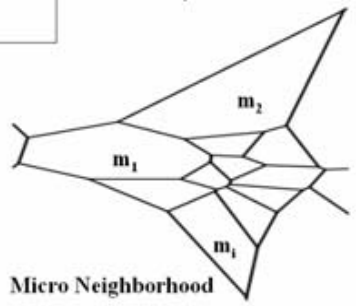
•Capture spatial heterogeneity by identifying **semantic relationships** between the non-spatial attributes and features in the micro neighborhoods

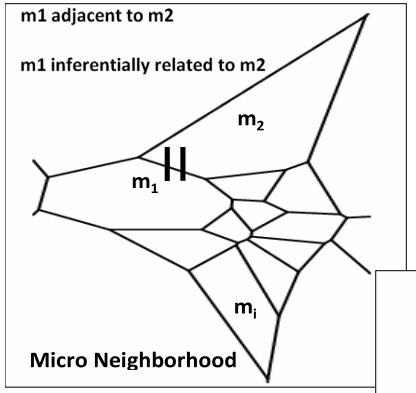
• A semantic relationship is defined using similarity coefficients such as Jaccard and Silhouette coefficients. If there is a spatial and semantic relationship between spatial objects the micro neighborhoods are merged to form macro **neighborhoods**

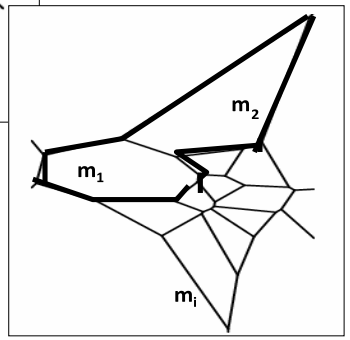
•The macro neighborhood captures both autocorrelation and heterogeneity in the neighborhood definitions by not only considering proximity but also the change in the localized features and attributes in the region.

- This neighborhood definition can be used as a precursor to the anomaly detection in reference to this neighborhood
- Similarly if an anomaly or an attack event is discovered in one part of this neighborhoods then chances are that the attack can propagate to other parts of the neighborhoods given the similarity and homogeneousness in this neighborhood



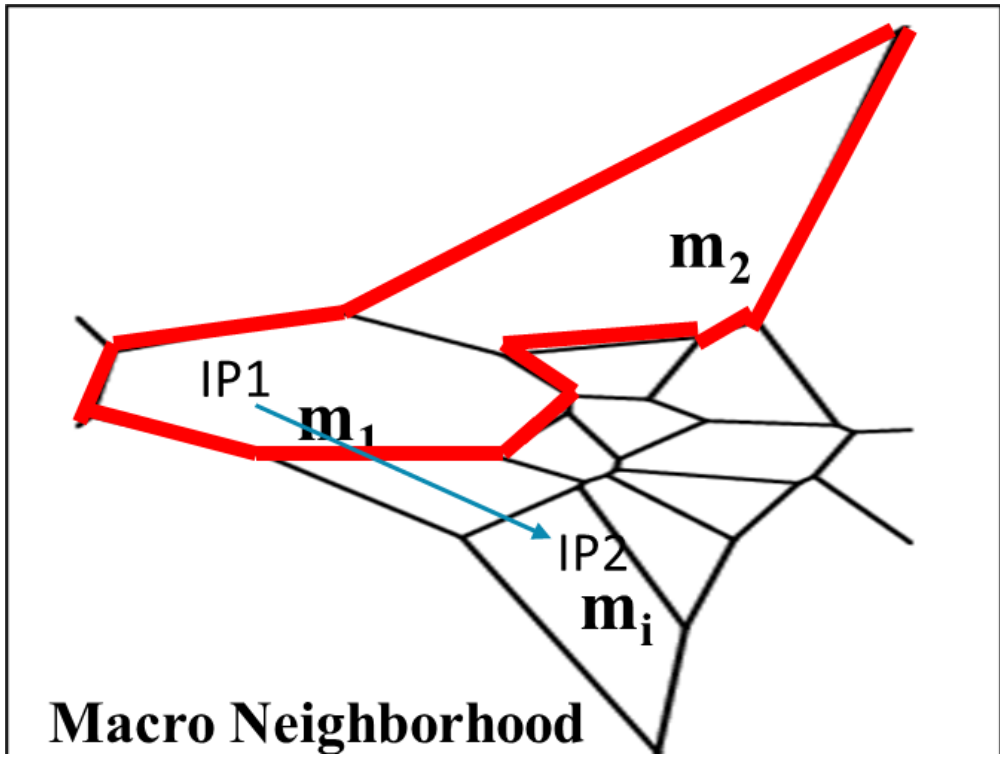






Spatial Neighborhood

- Neighborhood can be seen as an underlying characterization
- For example if two IP addresses are communicating with each other (e.g. IP1 sending a packet to IP2), then using geolocation (through databases such as Maxmind), these IP addresses can be located and overlaid on top of the neighborhood definitions
- For example if macro neighborhood m1 has a series of features indicating that it has known prior attacks originating, and is a source of geo-political unrest then this IP1 gets a certain reputation score and should be monitored carefully
- This will help network administrators go beyond the white and black listing of IP addresses and use geo located information sources as well.

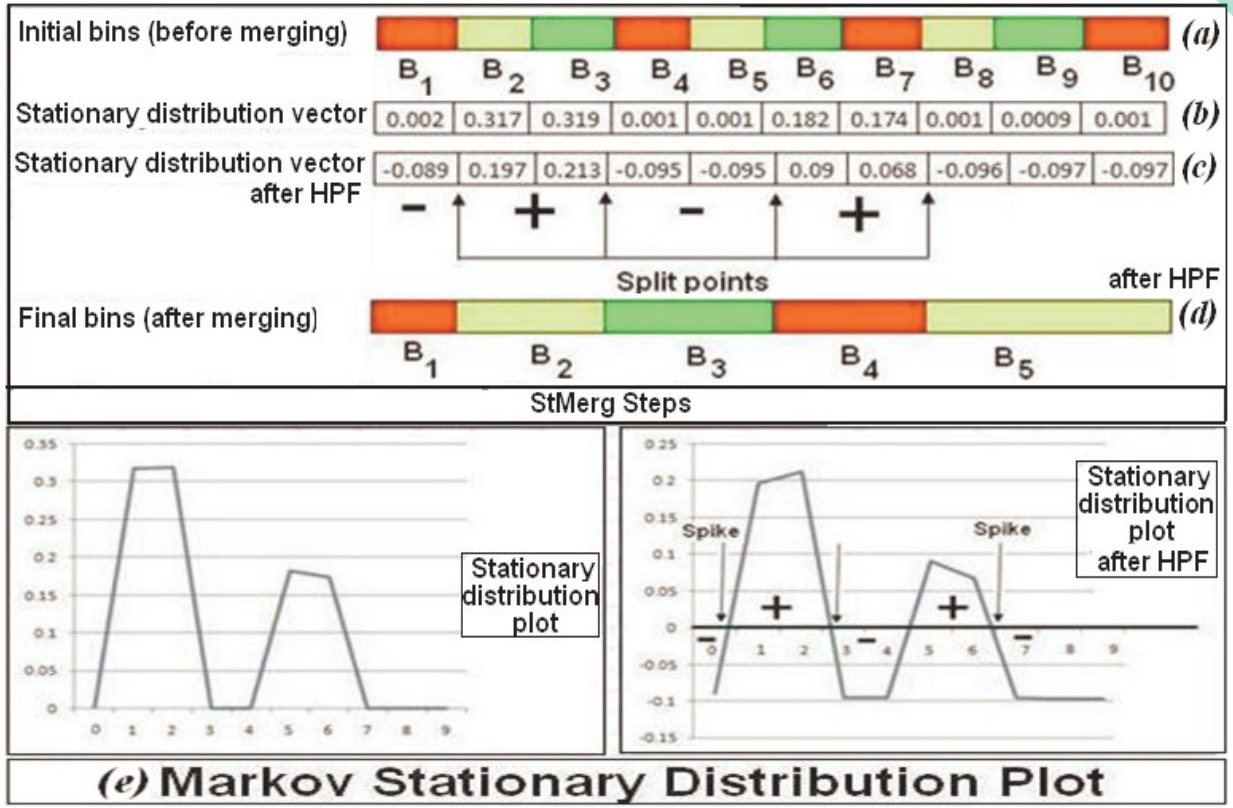


Temporal Anomaly Detection

- Change point detection addresses the discovery of time points at which the behavior of time series data changes
- Discretization has been extensively studied and many efficient solutions exist for simple discretization techniques including equal width or equiprobable discretization, Piecewise Aggregate Approximation (PAA)
- Time series *discords* are subsequences of a longer time series that are maximally different to all the rest of the temporal subsequences
- Most of these approaches consider autocorrelated segments as true segments for pattern discovery
- This may ignore heterogeneity in this process, possibly in the form of unequal width segments
- Most of these approaches do not quantify the unusualness of a patterns discovered in terms of how distinct is it from the data
- Similar to spatial behavior temporal autocorrelation and heterogeneity are a big challenge

Temporal Neighborhoods

- **Stationary Distribution based merging:** The data can be modeled as a Markov Model
- This approach starts with equal frequency bins as the states of a Markov model
- The similarity between the bins is computed using a distance measure *d* such as Kullback Leibler, Hellinger, and Mahalanobis, to name a few
- A transition matrix based on these similarities is generated for this Markov model and subsequently normalized to obtain a row-stochastic matrix
- In order to form an unequal depth discretization or in other words unequal sized temporal neighborhoods, an intuition is used that the adjacent bins in the transition matrix having high degree of probability of transition should be merged, in order to obtain unequal depth bins
- A Markov Stationary distribution based merging is used, which takes the transition matrix from the previous step and computes the stationary transition matrix by iteratively taking a self-product of the matrix until it converges or is near convergence such that every row entry in this converged matrix is same or nearly the same forming the stationary distribution vector
- The split points are identified by finding spikes in the stationary distribution vector plot. In order to detect these spikes in the stationary distribution vector
- Discrete Fourier Transform (DFT) and Inverse Discrete Fourier Transform (IDFT) are used with higher h coefficients as High Pass Filter (HPF)
- This temporal neighborhood can be used as a precursor to anomaly detection to identify anomalies in the temporal neighborhoods which are more homogenous for an accurate identification of unusual events
- The spikes in the distribution also may correspond to events of interest in the network traffic.



Cybersecurity through Spatiotemporal Analysis

Spatial Anomalies in Cybersecurity

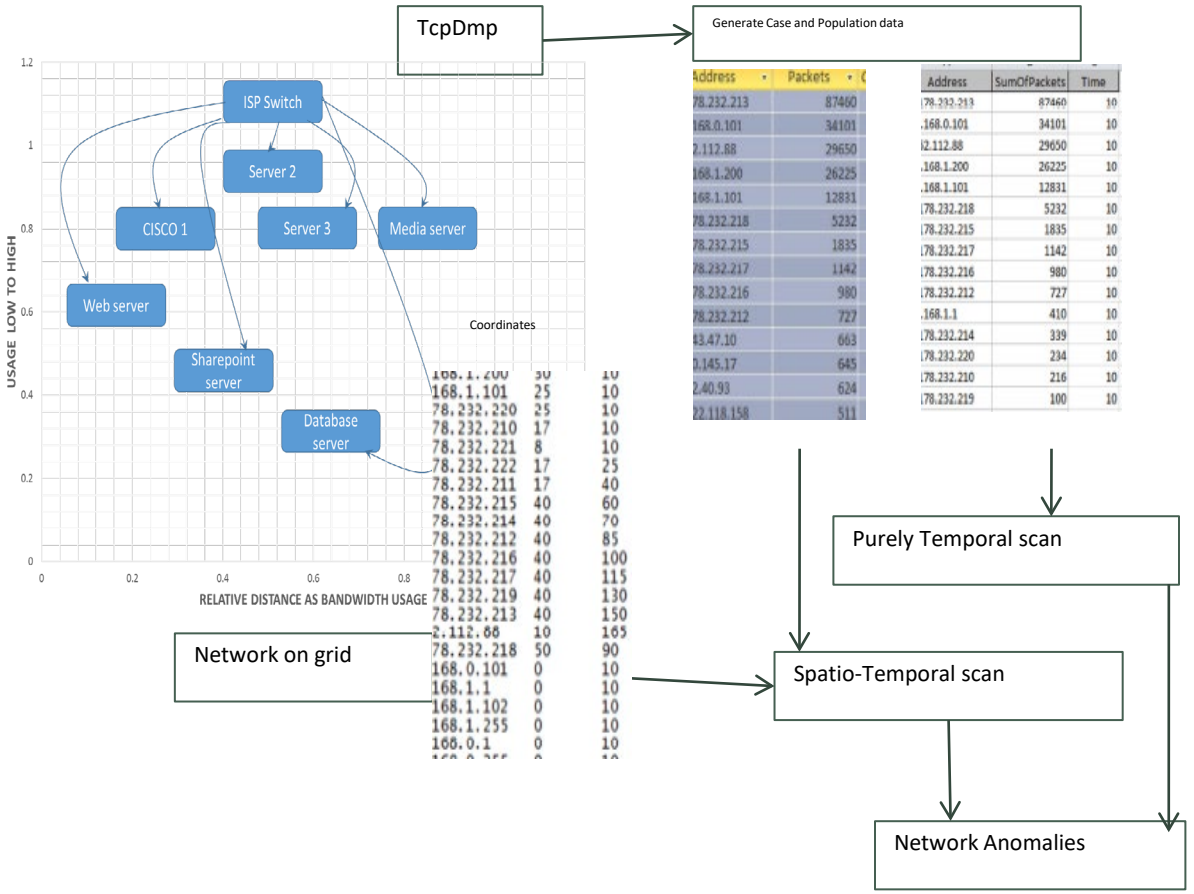
- While cyber threats that target networks are directed towards specific nodes during specific time periods, it has been noted that these nodes are not just randomly selected, but are usually located in specific geographical regions
- Kaspersky reported that Iran is the most common place where the Flame malware was discovered
- Cases were also reported in Israel, Palestine, Sudan, Syria, Lebanon, Saudi Arabia and Egypt
- Out of all computer systems that were affected worldwide by Stuxnet, 60% of them belonged to Iran A recent analysis indicates that most online schemes occur in Eastern Europe, East Asia, and Africa
- These schemes and phishing attacks are shifting to other countries like Canada, United States, France, Israel, Iran and Afghanistan
- The spatial element plays a key role in strengthening the cyber infrastructure, not only computer networks but also in power grids and industrial control systems
- While identifying locations of cyber attack victims is important in a power grid or computer network, it is also crucial to determine the location of those responsible for these attacks as it may have serious national security implications
- Most attacks are usually targeted towards specific countries, and for several different reasons
- A spatial aspect to cyber data analysis can be brought through:
  - Identification of spatial neighborhoods in the network structure to discover the impact of an event in space,
  - Studying influence of events in neighborhoods and
  - Studying correlation and associations of events over time

Neighborhood-Based Anomalies

- The neighborhood discovery can be used to describe influence flow among spatial objects
- In the network, vertices represent spatial objects, and edges represent the possible influence flow
- The weight on each edge indicates the hardship the influence faces to reach from one spatial object to another
- Such a weight takes into account the spatial distance and other factors such as barriers which can be network properties such as number of hops as the distance measure, geo political factors such as IP domain, country etc.
- So for instance, Mexico and United States, even though they are in proximity will have different rates of spread of an event due to differences in the cyber infrastructure
- Even though they are in proximity, their links weight will be less if we are studying event propagation
- This was made evident even within the United States when a major breach of Social Security Numbers occurred in South Carolina which did not affect the nearby or any other states
- Let us consider a network of nodes in a spatial neighborhood, where the spatial nodes are connected by edges, based on spatial relationships, with weights computed from the distance between the spatial and non-spatial attributes through similarity coefficients
- Given an event being studied, such as a specific attack if we want to study the propagation of the attack, influence distance can be used to quantify it
- The influence distance from one spatial object IP_p to object IP_q is the sum of the weights of the constituent edges of the shortest paths from IP_p to IP_q
- If IP_p and IP_q are unconnected then the influence distance is ∞. If IP_p to IP_q are the same then the distance is 0

Spatiotemporally Clustered Anomalies

- A network can also be seen as in a spatial layout to discover various types of clustered network anomalies
- A network can be laid out on a spatial grid based on relationships of network bandwidth usage and usage of the network
- Subsequently using any tcpdmp software the packet data can be extracted into population and case files where the population file has the overall network usage information extracted from historic data and case file has the current usage information
- Here network usage can be measured in terms of a variable of interest such as number of packets every minute or average packet size or total packet size
- Once the files are created from the raw data, spatial and temporal anomalies can be identified
- These anomalies are in the form of (a) cluster of servers which behave anomalously with respect to others as they are considered as geospatial points, (b) cluster of traffic points which are anomalous as a temporal cluster with respect to other neighboring servers



Temporal Behaviors in Evolving Networks

- Communication networks over a period of time can be studied by utilizing the temporal neighborhoods and help draw out several forms of communication patterns
- For example, one can identify if some nodes are popular on the network, either as a common source point to all other nodes or a common destination point from all other nodes on the network consistently
- One can also determine if the popularity of a node either increases or decreases over time
- Previous studies have discussed various properties in networks such as centrality of nodes in a network, the densification of a network, and the diameter of a network
- The mining of large networks can be performed with the aim of understanding the following:
    - Does the structure of the network change over time?
    - Can we define what changes occur as the network evolves?
    - Dan we identify when these changes in the network take place?
- By detecting changes in a temporally evolving network, one can drill down and identify the underlying cause of this change in the network, which could have come forth as a result of a threat to the network
- For example, the unexpected unavailability of a server system from the network could be due to an unknown denial of service attack

References

- D.A Griffith. Spatial Autocorrelation: A Primer. Association of American Geographers, 1987.
- R. Haining. Spatial Data Analysis: Theory and Practice. Cambridge University Press, Cambridge, UK, 2003.
- W.R. Tobler. A computer model simulation of urban growth in the detroit region. Economic Geogra- phy, 46(2):234–240, 1970
- Pusheng Zhang, Yan Huang, Shashi Shekhar, and Vipin Kumar. Correlation analysis of spatial time series datasets: A filter-and-refine approach. In the Proc. of the 7th Seventh Pacific-Asia Conference on Knowledge Discovery and Data Mining(PAKDD 2003), 2003.
- Raveh, A., & Tapiero, C. S. (1980). Periodicity, constancy, heterogeneity and the categories of qualitative time series. Ecology, 61(3), 715-719.
- Basawa, I. V., Billard, L., & Srinivasan, R. (1984). Large-sample tests of homogeneity for time series models. Biometrika, 71(1), 203-206.
- M. Ester, A. Frommelt, H.-P. Kriegel, and J. Sander. Algorithms for characterization and trend detec- tion in spatial databases. In 4th Int. Conf. on KDD, 1998.
- M. Ester, H. Kriegel, and J. Sander. Spatial data mining: A database approach. In 5th International Symposium on Advances in Spatial Databases, pages 47–66, London, UK, 1997. Springer-Verlag.
- I. Kang, T. Kim, and K. Li. A spatial data mining method by delaunay triangulation. In 5th ACM international workshop on Advances in Geographic Information Systems, pages 35–39, 1997.
- Y. Kou, C. Lu, and D. Chen. Spatial weighted outlier detection. In Proceedings of the Sixth SIAM International Conference on Data Mining, Bethesda, MD, USA, April 20-22, 2006. SIAM, 2006.
- S. Shekhar, C. Lu, and P. Zhang. Detecting graph-based spatial outliers: algorithms and applications (a summary of results). In 7th ACM international conference on Knowledge discovery and data mining, pages 371–376, 2001.
- Janeja, V. P., Adam, N. R., Atluri, V., & Vaidya, J. (2010). Spatial neighborhood based anomaly detection in sensor datasets. Data Mining and Knowledge Discovery, 20(2), 221-258.
- Aggarwal, C. C. (2017). Spatial Outlier Detection. In Outlier Analysis (pp. 345-368). Springer International Publishing.
- Janeja, V. P., & Atluri, V. (2009). Spatial outlier detection in heterogeneous neighborhoods. Intelligent Data Analysis, 13(1), 85-107.
- Ankerst, M., Breunig, M. M., Kriegel, H. P., & Sander, J. (1999, June). OPTICS: ordering points to identify the clustering structure. In ACM Sigmod record (Vol. 28, No. 2, pp. 49-60). ACM.
- Sander, J., Ester, M., Kriegel, H. P., & Xu, X. (1998). Density-based clustering in spatial databases: The algorithm gdbscan and its applications. Data mining and knowledge discovery, 2(2), 169-194.
- J. Glaz, J. Naus, and S. Wallenstein. Scan Statistics. Springer Verlag Series in Statistics, New York, 2001.
- J. Naus. The distribution of the size of the maximum cluster of points on the line. Journal of the American Statistical Association 60, pages 532–538, 1965.
- M. Kulldorff. A spatial scan statistic. Communications of Statistics - Theory Meth., 26(6):1481–1496, 1997.
- M. Kulldorff, W. Athas, E. Feuer, B. Miller, and C. Key. Evaluating cluster alarms: A space-time scan statistic and brain cancer in los alamos. American Journal of Public Health, 88(9):1377–1380, 1998.
- J. Besag and J. Newell. The detection of clusters in rare diseases. Journal of the Royal Statistical Society Series A, 154:143–155, 1991.
- S. Openshaw. A mark 1 geographical analysis machine for the automated analysis of point data sets. International Journal of GIS, 1(4):335–358, 1987.
- D. Neill, A. Moore, F. Pereira, and T. Mitchell. Detecting significant multidimensional spatial clusters. In Advances in Neural Information Processing Systems 17, pages 969–976, Cambridge, MA, 2005. MIT Press.
- T. Tango and K. Takahashi. A flexibly shaped spatial scan statistic for detecting clusters. International Journal of Health Geographics, 4(11), 2005.
- Vijay S. Iyengar. On detecting space-time clusters. In KDD '04: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, pages 587–592, New York, NY, USA, 2004. ACM Press.
- L. Duczmal and A. Renato. A simulated annealing strategy for the detection of arbitrarily shaped spatial clusters. Comput. Statist. Data Anal. To appear, 2003.

References

- V.P. Janeja and V. Atluri. ls3: A linear semantic scan statistic technique for detecting anomalous windows. In ACM Symposium on Applied Computing, 2005.
- V.P. Janeja and V. Atluri. FS3: A random walk based free-form spatial scan statistic for anomalous window detection. In Fifth IEEE International Conference on Data Mining(ICDM'05), pages 661– 664. IEEE Computer Society, 2005.
- Janeja, V. P., & Atluri, V. (2008). Random walks to identify anomalous free-form spatial scan windows. IEEE Transactions on Knowledge and Data Engineering, 20(10), 1378-1392.
- Shi, L., & Janeja, V. P. (2009, June). Anomalous window discovery through scan statistics for linear intersecting paths (SSLIP). In Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 767-776). ACM.
- H. J. Miller. Tobler's first law and spatial analysis. Annals of the Assoc. of American Geographers, 94(2):284289, 2004.
- O. Sugiura. Testing change-points with linear trend, 1994.
- Kenji Yamanishi and Jun ichi Takeuchi. A unifying framework for detecting outliers and change points from non-stationary time series data. In KDD '02: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, pages 676–681, New York, NY, USA, 2002. ACM.
- Valery Guralnik and Jaideep Srivastava. Event detection from time series data. In KDD '99: Proceed- ings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining, pages 33–42, New York, NY, USA, 1999. ACM.
- Duchene. F., Garbayl. C., and Rialle. V. Mining heterogeneous multivariate time-series for learning meaningful patterns: Application to home health telecare, 2004.
- Jessica Lin, Eamonn Keogh, Stefano Lonardi, and Bill Chiu. A symbolic representation of time series, with implications for streaming algorithms. In DMKD '03: Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery, pages 2–11, New York, NY, USA, 2003. ACM.
- O. Tim, L. Firoiu, and P. Cohen. Clustering time series with hidden markov models and dynamic time warping. In Presented at IJCAI-99 Workshop on Sequence Learning., 1999.
- Jamal Ameen and Rawshan Basha. Mining time series for identifying unusual sub-sequences with ap- plications. In ICICIC '06: Proceedings of the First International Conference on Innovative Computing, Information and Control, pages 574–577, Washington, DC, USA, 2006. IEEE Computer Society.
- Keogh. E., Lin. J., and Fu. A. Hot sax: efficiently finding the most unusual time series subsequence. In Fifth IEEE International Conference on Data Mining. IEEE, 2005.
- Dragomir Yankov, Eamonn Keogh, and Umaa Rebbapragada. Disk aware discord discovery: Finding unusual time series in terabyte sized datasets. In Seventh IEEE International Conference on Data Mining. IEEE, 2007.
- Li Wei, Eamonn Keogh, and Xiaopeng Xi. Saxually explicit images: Finding unusual shapes. In ICDM '06: Proceedings of the Sixth International Conference on Data Mining, pages 711–720, Washington, DC, USA, 2006. IEEE Computer Society.
- Tamas Abraham and John F. Roddick. Survey of spatio-temporal databases. GeoInformatica, 3(1):61– 99, 1999.
- John F. Roddick and Kathleen Hornsby, editors. Temporal, Spatial, and Spatio-Temporal Data Mining, First International Workshop TSDM 2000 Lyon, France, September 12, 2000, Revised Papers, volume 2007 of Lecture Notes in Computer Science. Springer, 2001.
- John F. Roddick, Kathleen Hornsby, and Myra Spiliopoulou. An updated bibliography of temporal, spatial, and spatio-temporal data mining research. In TSDM '00: Proceedings of the First International Workshop on Temporal, Spatial, and Spatio-Temporal Data Mining-Revised Papers, pages 147–164, London, UK, 2001. Springer-Verlag.
- John F. Roddick and Myra Spiliopoulou. A bibliography of temporal, spatial and spatio-temporal data mining research. SIGKDD Explor. Newsl., 1(1):34–38, 1999.
- S. Dey, V. P. Janeja, and A. Gangopadhyay. Temporal neighborhood discovery through unequal depth binning. In IEEE International Conference on Data Mining(ICDM'09), 2009
- Sandipan Dey, Vandana Pursnani Janeja, Aryya Gangopadhyay:
- Discovery of temporal neighborhoods through discretization methods. Intell. Data Anal. 18(4): 609-636 (2014)
- E.O. Brigham. The Fast Fourier Transform. New York: Prentice-Hall, 2002.

References

- C. Burns. Stuxnet virus origin confirmed: Usa and isreali governments., 6/1/12 2012. http://www.slashgear.com/stuxnet-virus-origin-confirmed-usa-and-isreali-governments-01231244/.
- J. Halliday. Stuxnet worm is the 'work of a national government agency'., 9/24/2010 2010. http://www.guardian.co.uk/technology/2010/sep/24/stuxnet-worm-national-agency.
- J. Bronskill. Govt fears canada becoming host country for cyber-attacker., 11/9/2012 2012. http://www.ctvnews.ca/canada/gov-t-fears-canada-becoming-host-country-for-cyber-attackers- 1.1032092#ixzz2BluFFbM7.
- M. P. McGuire, V.P. Janeja, and A. Gangopadhyay. Spatiotemporal neighborhood discovery for sensor data. In Proceedings of the 2nd International Workshop on Knowledge Discovery from Sensor Data (Sensor-KDD 2007), held in conjunction with the 14th International Conference on Knowledge Discovery and Data Mining (ACM SIG-KDD 2008), August 2008.
- M. P. McGuire, V.P. Janeja, and A. Gangopadhyay. Mining sensor datasets with spatio-temporal neighborhoods. Journal of Spatial Information Science (JOSIS), 2012. Accepted 2012.
- Yanan Sun, Vandana Pursnani Janeja, Michael P. McGuire, and Aryya Gangopadhyay. Tnet: Tensor- based neighborhood discovery in traffic networks. In ICDE Workshops, pages 331–336, 2012.
- U. Kang, C. Tsourakakis, and C. Faloutsos. Pegasus: A peta-scale graph mining system - implemen- tation and observations. In ICDM, 2009.
- U. Kang, C. Tsourakakis, A. Appel, C. Faloutsos, and J. Leskovec. Radius plots for mining tera-byte scale graphs: Algorithms, patterns, and observations. In SIAM International Conference on Data Mining (SDM)., 2010.
- J. Leskovec. Dynamics of large networks., 2008.
- J. Leskovec, D. Chakrabarti, J. Kleinberg, and C. Faloutsos. Realistic, mathematically tractable graph generation and evolution, using kronecker multiplication. In European Conference on Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD), 2005.
- J. Leskovec and C. Faloutsos. Scalable modeling of real graphs using kronecker multiplication. In International Conference on Machine Learning (ICML), 2007.
- J. Leskovec, J. Kleinberg, and C. Faloutsos. Graphs over time: Densification laws, shrinking diame- ters and possible explanations. In ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2005.
- J. Leskovec, J. Kleinberg, and C. Faloutsos. Graph evolution: Densification and shrinking diameters. In ACM Transactions on Knowledge Discovery from Data (TKDD), volume 1, 2007.
- Ying Ding, Erjia Yan, Arthur Frazho, and James Caverlee. Pagerank for ranking authors in co-citation networks. J. Am. Soc. Inf. Sci. Technol., 60(11):2229–2243, November 2009
- G. Phillips, S. Shenker, and H. Tangmunarunkit. Scaling of multicast trees: Comments on the chuang- sirbu scaling law. In SIGCOMM '99: Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication., 1999.
- A. Fabrikant, E. Koutsoupias, and C. H. Papadimitriou. Heuristically optimized trade-offs: A new paradigm for power laws in the Internet, volume 2380 of Automata, Languages and Programming, page 781. Springer, 2002.
- M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. In SIGCOMM, 1999.
- Gu, G., Zhang, J., & Lee, W. (2008, February). BotSniffer: Detecting Botnet Command and Control Channels in Network Traffic. In NDSS (Vol. 8, pp. 1-18).
- Bradley Isbell, Hamad Alsaleh,Thuy Lam, V. Janeja, Analysis of SSH Authentication Attacks, Technical report, UMBC, December 2016.
- Keim, D. A., Mansmann, F., & Schreck, T. (2005). Analyzing electronic mail using temporal, spatial, and content-based visualization techniques. In Informatik 2005-Informatik live! (pp. 434-438).
- Keim, D. A., Mansmann, F., Panse, C., Schneidewind, J., & Sips, M. (2005). Mail explorer-spatial and temporal exploration of electronic mail. In EuroVis (pp. 247-254).
- Gu, G., Zhang, J., & Lee, W. (2008, February). BotSniffer: Detecting Botnet Command and Control Channels in Network Traffic. In NDSS (Vol. 8, pp. 1-18).