Electronics and Telecommunications Research Institute

수행과제 소개

16 September 2017

Sung-Soo Kim

sungsoo@etri.re.kr

Smart Data Research Group

ETRI

Outline

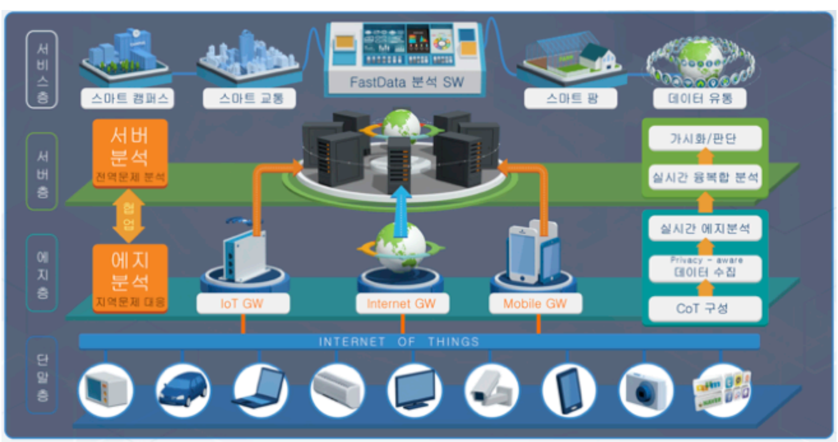
- 수행과제 소개
  - CoT 과제
  - 교통 과제

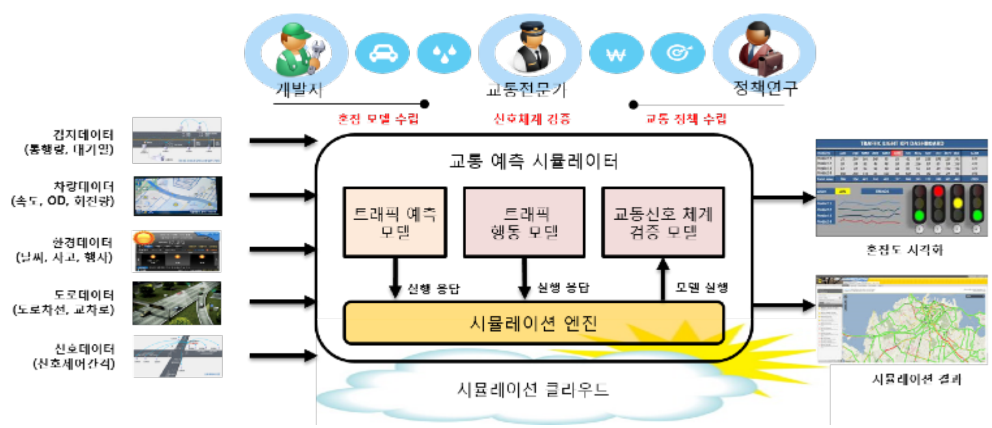
- Smart Data and Big Data Analytics
  - *Predictive Policing*

수행 과제

- [CoT 과제]
  - CoT(Cloud of Things) 환경에서 실시간 반응성 향상을 위한 계층적 데이터 스트림 분석 SW 기술 개발

- [교통 과제]
  - 도시 교통 문제 개선을 위한 클라우드 기반 트래픽 예측 시뮬레이션 SW 기술 개발





Electronics and Telecommunications Research Institute

CoT (Cloud of Things) Project

연구배경

"Fast Data: The next step after big data."
(Infoworld, Jun. 2014)

"Fast Data is Breaking Your Architecture"

※ Fast Data? "Incredible Speed"로 생성되는 데이터
(센서/IoT 데이터, 네트워크, 공정과정, 뉴스, SNS, 로그, 금융 등)

Based on DB

인메모리 NewSQL 데이터베이스

FastData의 주요기술

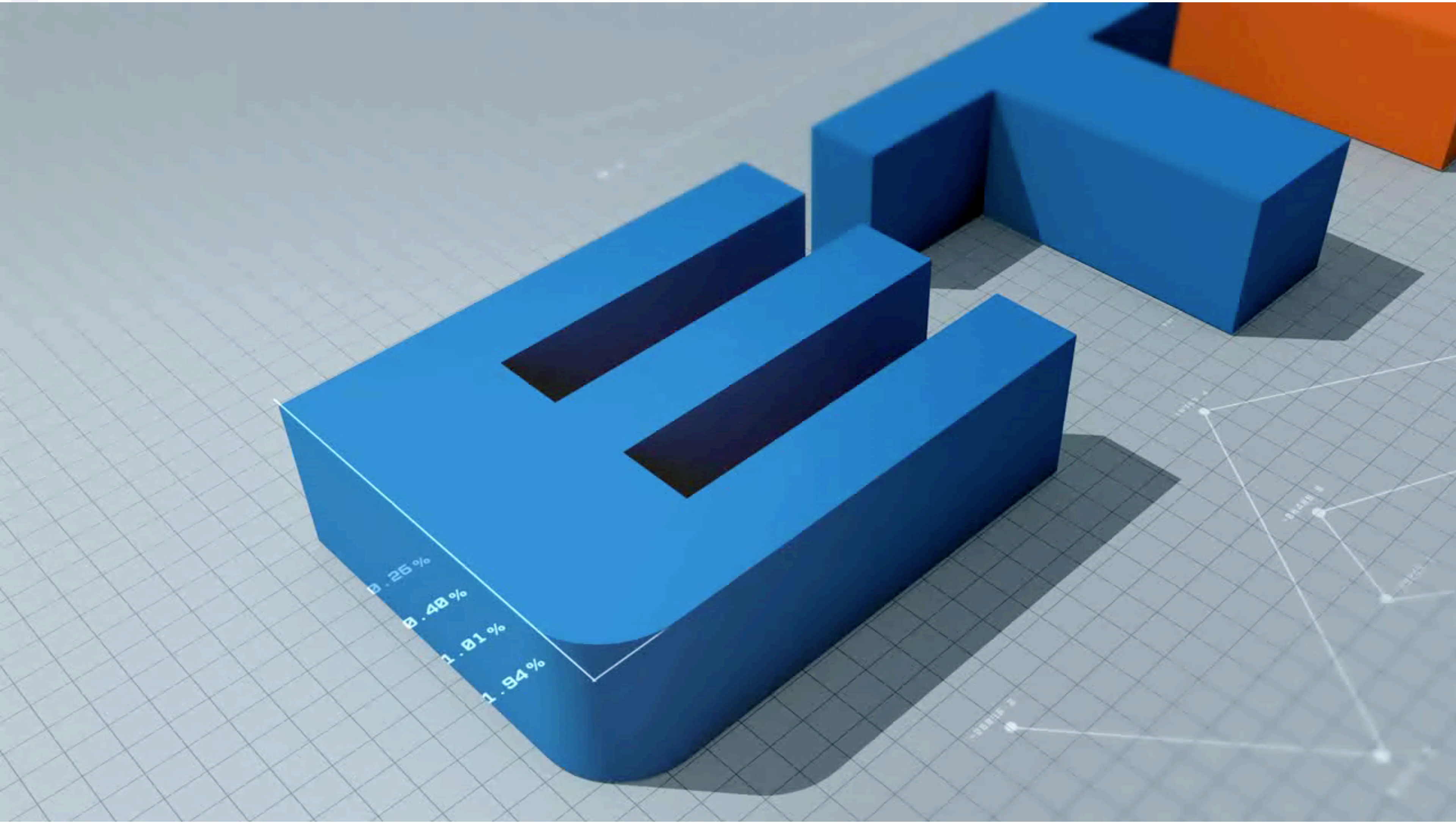
New SW Architecture

Fast Reaction을 위한 스트림 데이터 처리 인프라 기술

필터링, 상호연관 (Filtering,..) → 모델링 (Modeling) → 분석 & 가치 (Analytics) → 조치 (Decision)

"Data in Rest → Fast Data in Motion"

과제 배경 및 개념



연구배경

Fast Data 를 효율적으로 처리하는 아키텍처는?

IoT 환경의 수많은 Things

초연결 사회에서 Things을 좀 더 쉽게 표현하는 방법?

용도, 지역별 논리적인 그룹 구성 → CoT (Cloud of Things)

※ CoT(Cloud of Things): 사물의 연결성, 데이터 저장 및 분석을 위해 가상화되어 구성된 사물정보의 그룹

분주한 서버 자원 vs. 유휴자원을 가진 에지

에지의 유휴 자원을 활용하여 서버의 트래픽을 줄이자.

에지의 자원을 활용하는 에지 컴퓨팅

※ 에지(Edge) : 최종 단말과 지역적으로 근접한 각종 게이트웨이
(예) IoT GW, Phone, Network GW, Local Server, Home Server 등

빠른 처리를 요구하는 사회 현안들

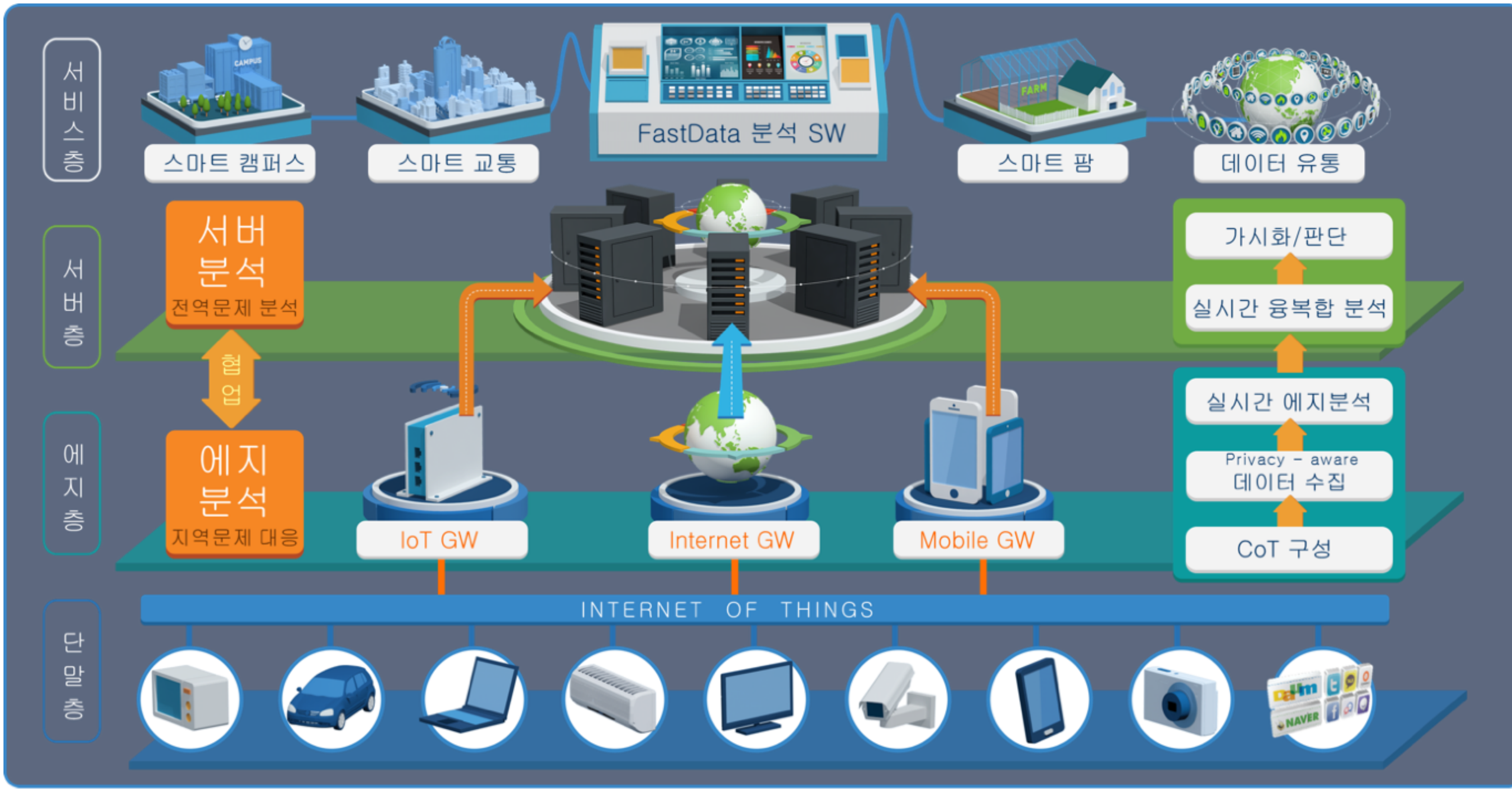
중앙서버의 판단을 기다리지 않고, 지역적인 신속 반응 방법은?

빠른 응답과 서버 단절 시에도 가능한 에지 상의 처리/분석

※ 시스코의 Fog Computing, CMU의 Cloudlet 등

연구목표

- CoT 환경에서 발생하는 FastData에 대하여
  - ① 빠르게 대응할 수 있는 에지 분석 및 CoT 구성 기술과
  - ② 전역적인 문제해결을 위한 에지-서버 간 지능적인 계층적 협업분석 기술 확보



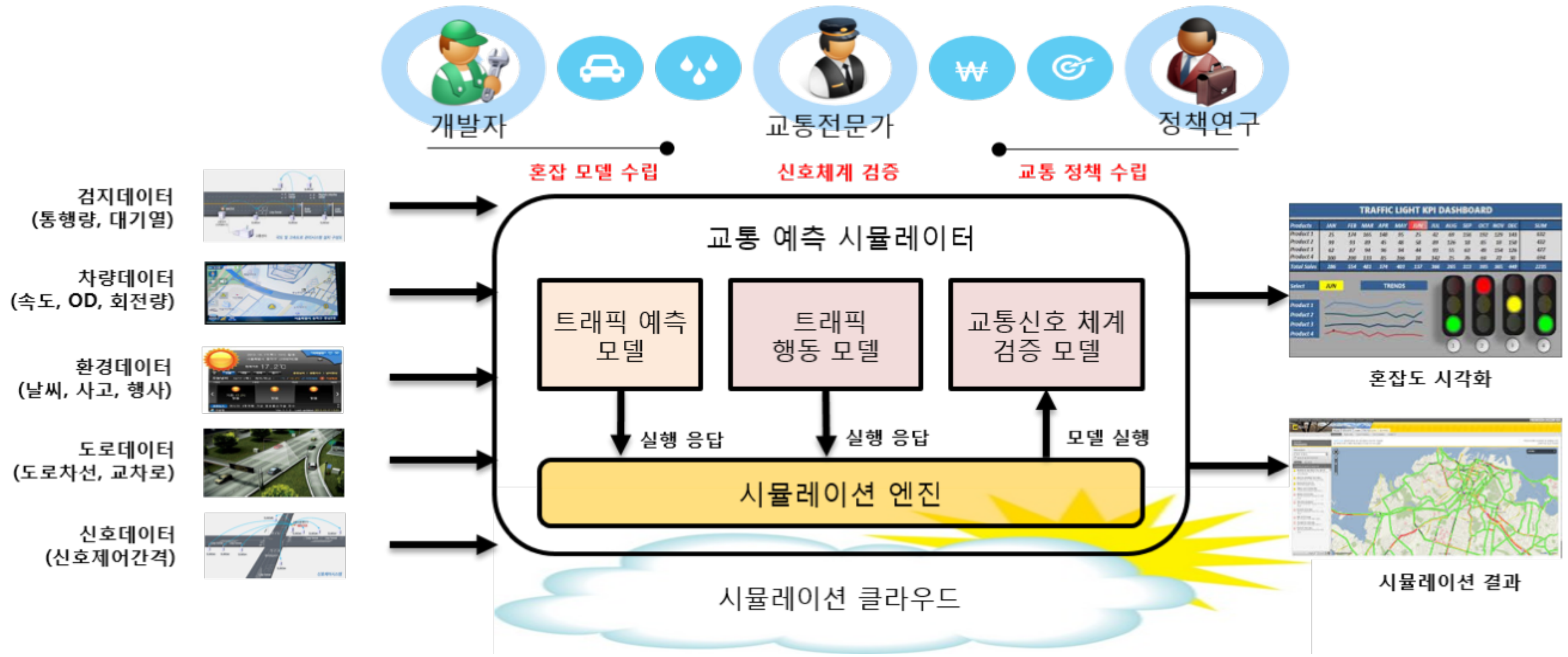
Electronics and Telecommunications Research Institute

교통과제

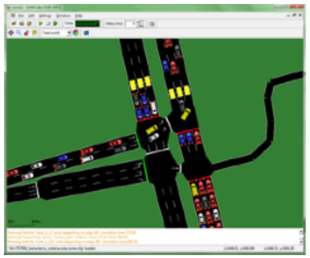
도시 교통 문제 개선을 위한 클라우드 기반 트래픽 예측 시뮬레이션 SW 기술 개발

연구목표

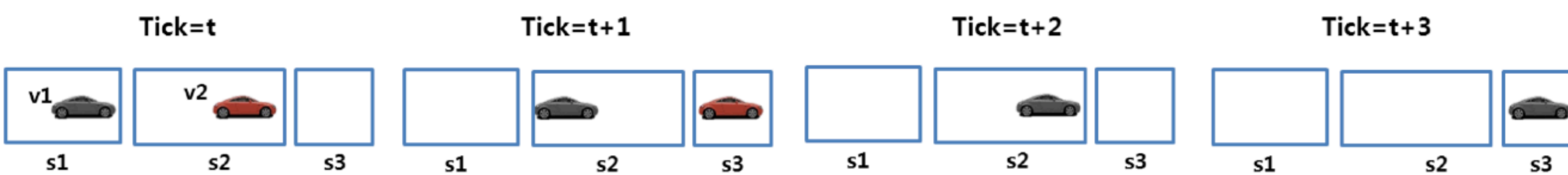
최종목표	도심의 교통 혼잡 문제를 완화시키기 위해, 공공/민간 교통정보 데이터를 통합 활용하여 도심의 **교통 상황을 예측**하고 **클라우드 인프라 확장성**을 갖춘 아키텍처를 도입하여 동적 규모로 **교통 신호체계 개선 방법** 및 **교통 정책을 사전 검증**할 수 있는 **트래픽 예측 시뮬레이션 기술 개발**



Microscopic vs. Mesocopic Traffic Simulation



Scale	Input	Output
Micro	Road network Traffic demand (Traffic Light System)	매 tick당 개별 차량의 위치
Meso		개별 차량의 도로 진입/이탈 시각



$Time$	$[(veh, position), ..]$
t	$[(v_1, s_1, 10), (v_2, s_2, 20)]$
$t+1$	$[(v_1, s_2, 0), (v_2, s_3, 0)]$
$t+2$	$[(v_1, s_2, 20), (v_2, s_3, 0)]$
$t+3$	$[(v_1, s_3, 0)]$

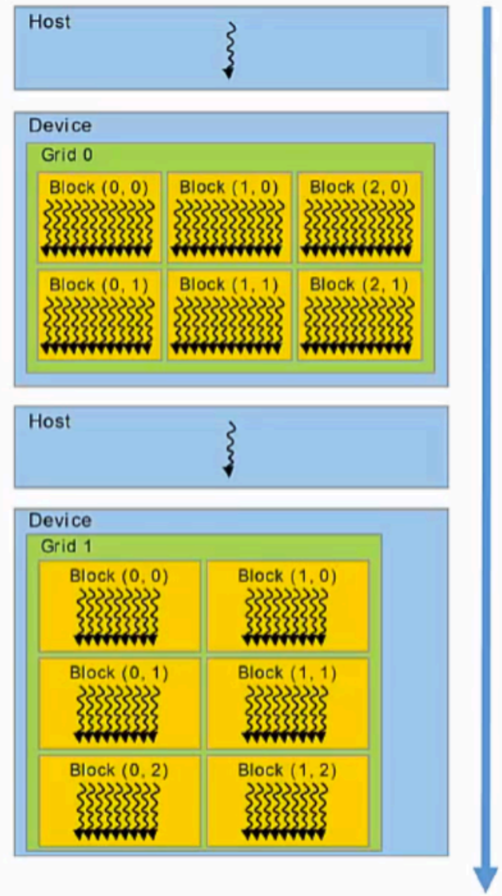
Seg	$[(veh, t_{enter}, t_{leave}), \dots]$
s_1	$[(v_1, t, t+1)]$
s_2	$[(v_1, t+1, t+3), (v_2, -, t+1]$
s_3	$[(v_1 t+3, -), (v_2 t+1, t+2)]$

CUDA Basics

Program Flow

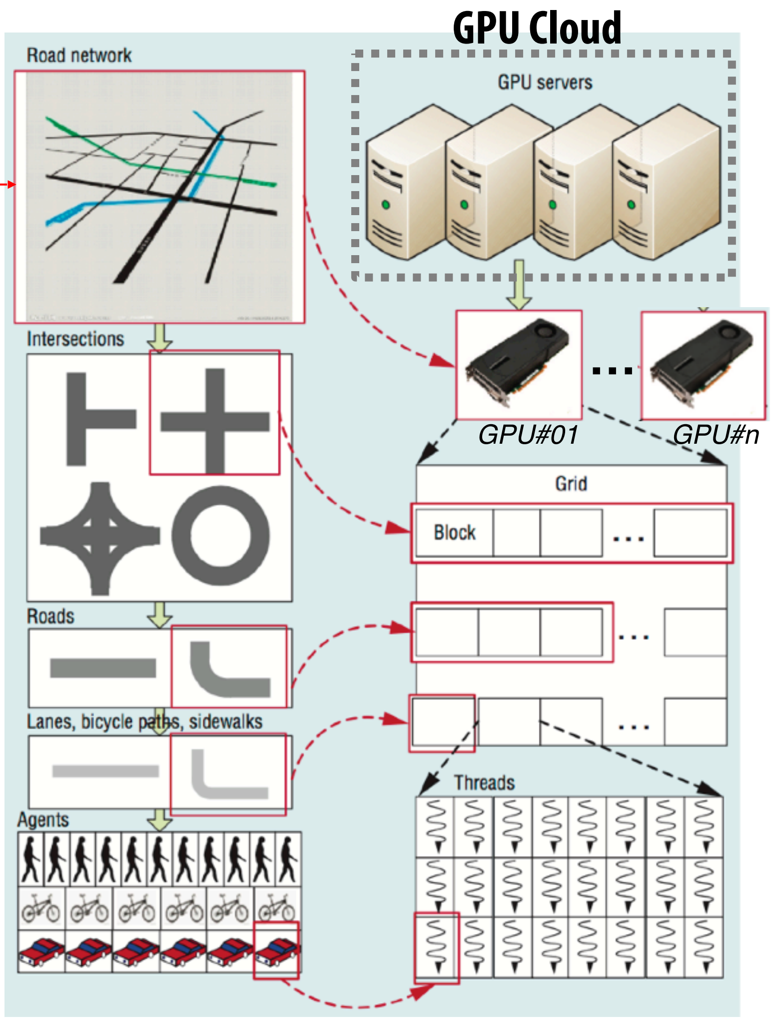
- Host Code
  - Serial
  - Runs on CPU
  - Launches Kernel's
- Device Code
  - Parallel
  - Runs on GPU

```
int main( void ) { // Host Code
   // Do sequential stuff

   // Launch Kernel
   kernel_0<<< grid_sz0,blk_sz0 >>>(...);

   // Do more sequential stuff

   // Launch Kernel
   kernel_1<<< grid_sz1,blk_sz1 >>>(...);

    return 0;
}
```



GPU기반 분산/병렬 시뮬레이션 기술

대규모 교통 시뮬레이션



Electronics and Telecommunications
Research Institute

Predictive Policing 101

Preventing Crime with Data and Analytics

16 September 2017

Sung-Soo Kim

sungsoo@etri.re.kr

Smart Data Research Group

ETRI

References

1. Craig D. Uchida et. al, ***Data-Driven Crime Prevention - New Tools for Community Involvement and Crime Control***, Justice & Security Strategies, Inc., 2013.
2. Walter L. Perry et.al, ***Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations***, Research Report of RAND Safety and Justice Program, 2013.
3. Jennifer Bachner, ***Predictive Policing; Preventing Crime with Data and Analytics***, Improving Performance Series, IBM Center for The Business of Government, 2013.
4. Ryan Gale, ***An Application of Risk Terrain Modeling to Residential Burglary***, TCNJ Journal, 2013.
5. James Byrne, ***A History of Crime Analysis***, Technology and the Criminal Justice Program (44.203 Course Note), UMass Lowell, Fall, 2012.
6. Mohler, G.O., M.B. Short, P.J. Brantingham, F.P. Schoenberg, and G.E. Tita, ***Self-exciting point process modeling of crime***. *Journal of the American Statistical Association* 106(493):100-108, 2011.
7. Erik Lewis *et. al*, ***Self-exciting point process models of civilian deaths in Iraq***, *Security Journal*, 2011.
8. Leslie Kennedy, Joel Caplan and Eric Piza, ***Risk Clusters, Hotspots, and Spatial Intelligence: Risk Terrain Modeling as an Algorithm for Police Resource Allocation Strategies***, *Journal of Quantitative Criminology* 27, pp. 339–362, 2011.
9. Joel Caplan, Leslie Kennedy, and Joel Miller, ***Risk Terrain Modeling: Brokering Criminological Theory and GIS Methods for Crime Forecasting***, *Justice Quarterly* 28:2, pp. 360–381, 2011.
10. Graham Farrell and William Sousa, ***Repeat Victimization and Hot Spots: The Overlap and its Implications for Crime Control and Problem-Oriented Policing***, Crime Prevention Studies 12, pp. 221–240, 2001.
11. Mike Egesdal *et. al*, ***Statistical and Stochastic Modeling of Gang Rivalries in Los Angeles***, *SIAM Journal*, 2010.
12. AUSTIN C. ALLEMAN, ***GEOGRAPHIC PROFILING THROUGH SIX-DIMENSIONAL NONPARAMETRIC DENSITY ESTIMATION***, *Technical Report*, DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, SANTA CLARA UNIVERSITY, 2012.
13. P Jeffrey Brantingham, ***Prey selection among Los Angeles car thieves***, *Crime Science* 2013.
14. George Mohler, ***MODELING AND ESTIMATION OF MULTI-SOURCE CLUSTERING IN CRIME AND SECURITY DATA***, *Annals of Applied Statistics*, 2013.

Depth of Knowledge (DoK) Level of the Seminar

Outline

- Why is the Predictive Policing Important?
- Background
- Mathematical Frameworks
- Data Used in Predictive Policing
- Predictive Methodologies for Predictive Policing
- Summary

Why is the Predictive Policing Important?

Why is the Predictive Policing Important?

- What is Crime Analysis?
  - Administrative, Tactical and Strategic Problem Solving
  - Forecasting and predictive analysis

August Vollmer
Source: LAPD Web Site
1909

- Principles of *problem-oriented* policing
- *Evidence-based* policing
- *Real-time* crime analysis
- *Intelligence-led* policing and other proven policing models

Orlando W. Wilson
1950

ETRI Electronics and Telecommunications Research Institute

History

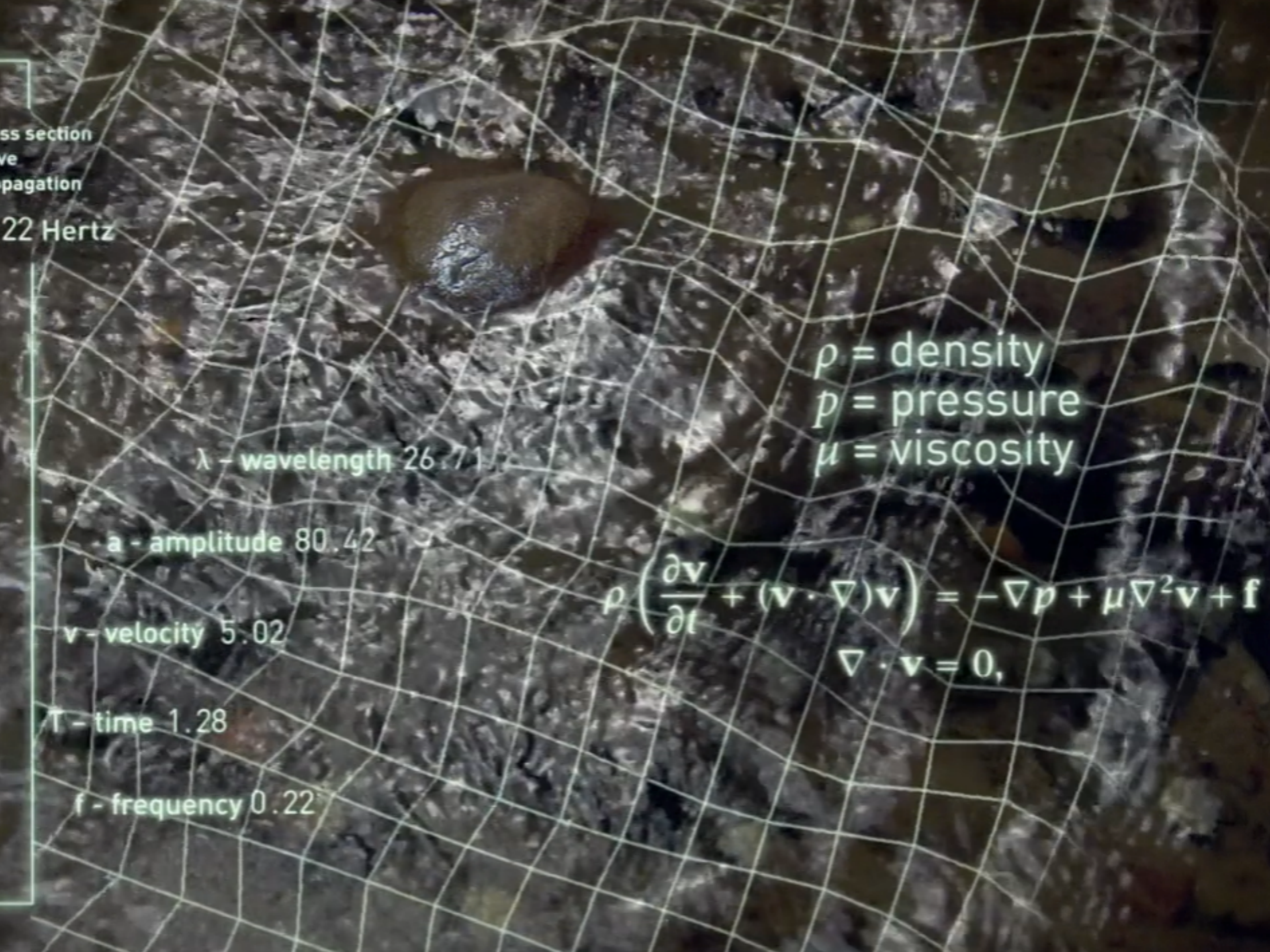
- *Origin of Crime Analysis*
  - Crime Mapping: 1829, Italian geographer and French Statistician
  - Crime Analysis: 1842, London's Metropolitan Police

- *First Data Miner*
  - John Graunt [Amateur Scientist]
  - Bill of Mortality (1st: 21 December 1592)
    - essentially *random* set of information
    - the study of the patterns causes and effects of *disease*
    - *Plague*
      - caused by person-to-person contact
      - tended to increase during the first year of the reign of a new king
    - gold mining


A generall Bill for this present year, ending the 19 of December 1665. according to the Report made to the KINGS most Excellent Majesty.
By the Company of Parish Clerks of London, &c.

The Total of all the Christnings — 9967
The Total of all the Burials this year — 97306
Whereof, of the Plague — 68596

The Diseases and Casualties this year.


ss section
ve
pagation
22 Hertz
λ - wavelength 26.71
a - amplitude 80.42
v - velocity 5.02
T - time 1.28
f - frequency 0.22
ρ = density
p = pressure
μ = viscosity
$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v}\right) = -\nabla p + \mu\nabla^2\mathbf{v} + \mathbf{f}$
$\nabla\cdot\mathbf{v} = 0,$

Clustering Patterns

- Loma Prieta Earthquake (1989)
  - Earthquake analysis
  - Clustering patterns
- MASC Project
  - Clustering patterns are also seen in *crime data*
  - Future events *nearby* in *space* and *time* after shocks of crime
  - Finding and clustering the *patterns* from chaos human behaviours
  - *Self-exciting point* process model *similar* to that used in earthquake analysis
  - Pilot Project : Los Angeles
    - Data: 13 million past crimes
    - Duration: past 80 years

UC MASC PROJECT
MATHEMATICAL AND SIMULATION MODELING OF CRIME

Mathematical and Simulation Modeling of Crime

UC MaSC Project
Mathematical and Simulation Modeling of Crime

Jeff Brantingham
UCLA Anthropology

George Mohler
Santa Clara Mathematics

George Tita
UCI Criminology, Law and Society

Mathematical Framework [7]

- **The Hawkes Process Model (1974)**
  - any given event, or collection events can be causally linked to a *background Poisson process* and *foreground self-exciting process*.

Self-excitation

the distribution of crimes following an initial event

$$\lambda(t) = \mu + k_0 \sum_{t_k < t} g(t - t_k)$$

background rate of events
stationary Poisson process

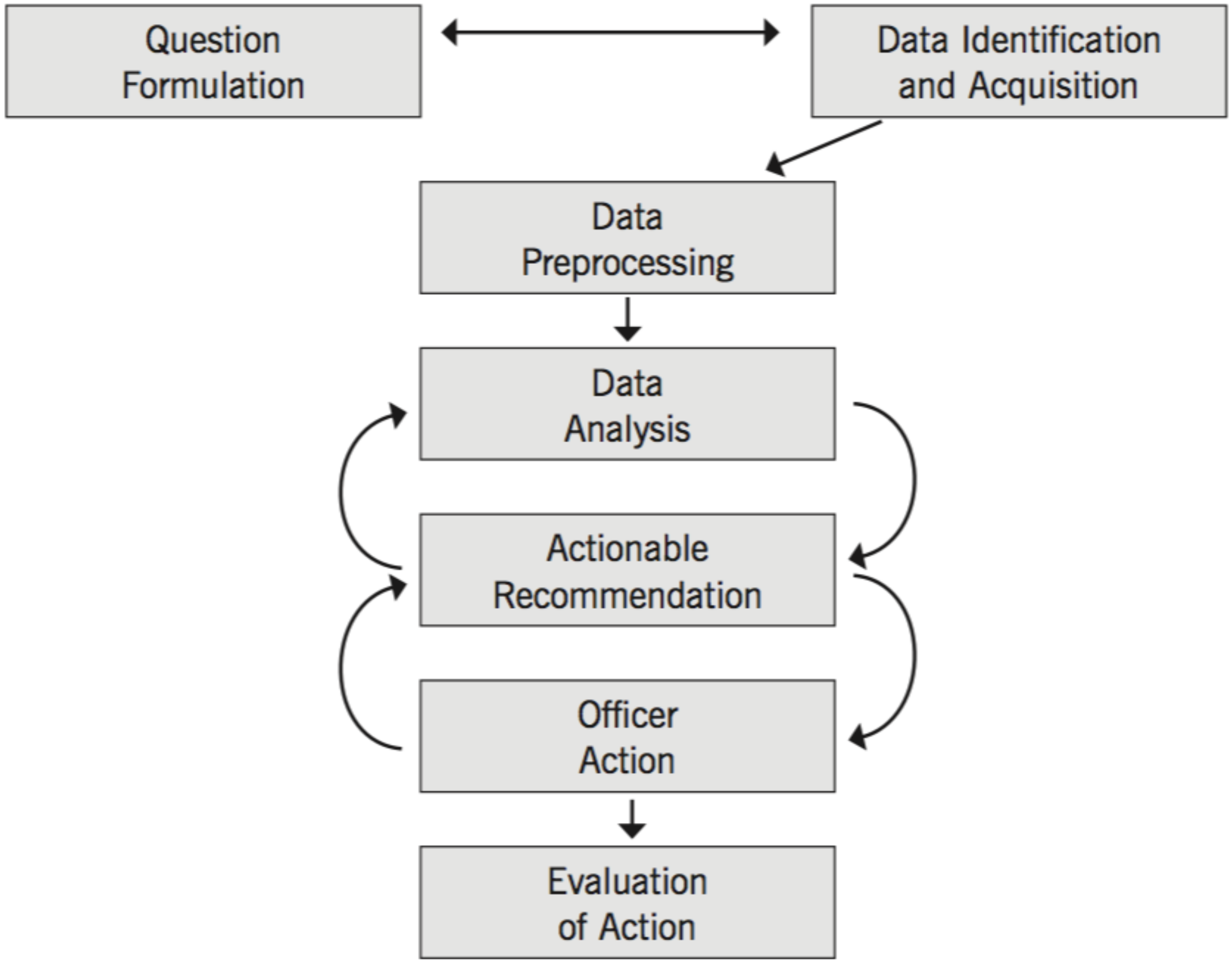
the density of prior events necessary to trigger excitation

how much excitation is generated by a collection of prior events

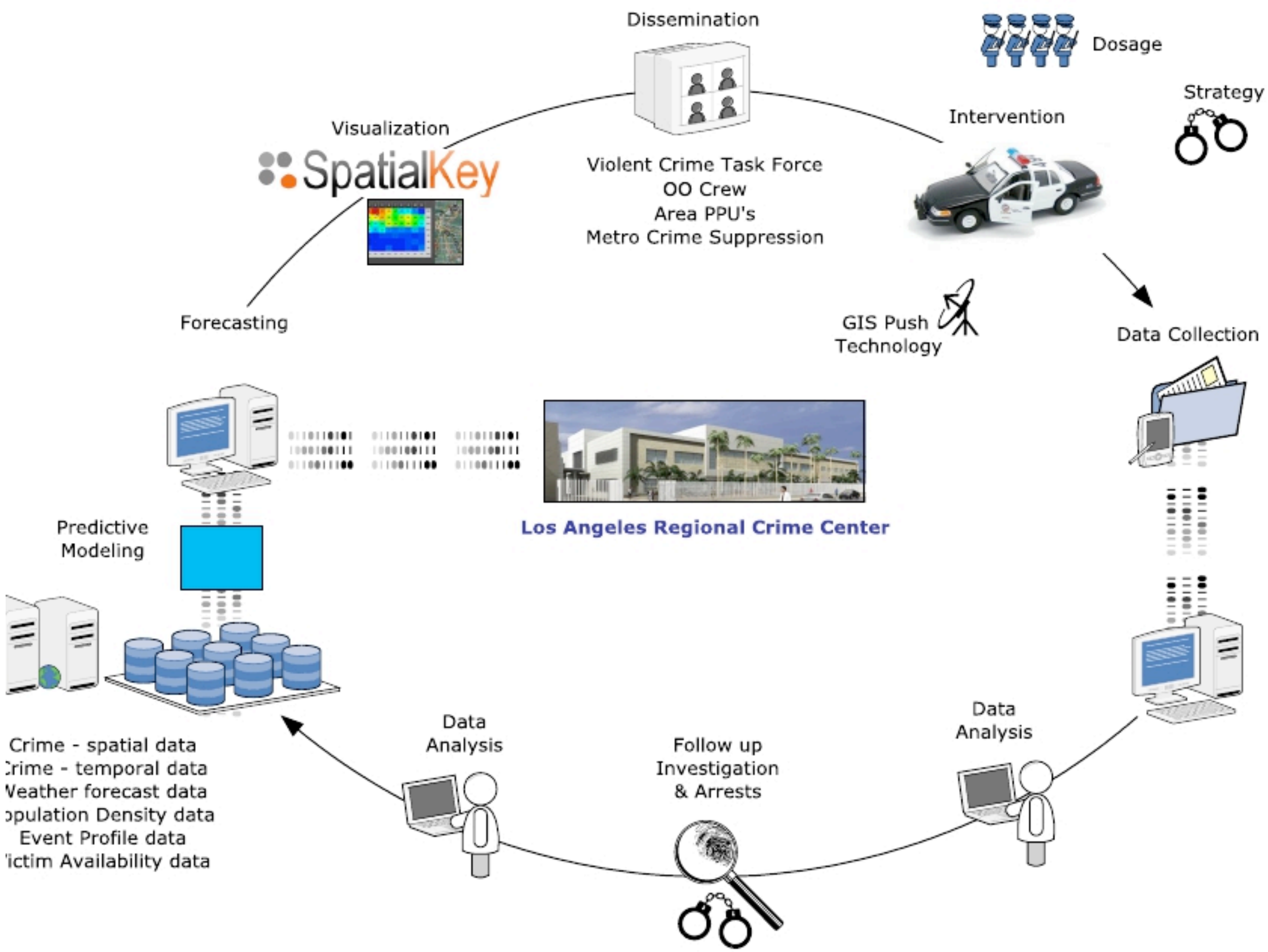
[7] Erik Lewis *et. al*, ***Self-exciting point process models of civilian deaths in Iraq***, *Security Journal*, 2011.


BURGLARY
CASES

Operational Challenges of Predictive Policing



Case Study: Los Angeles Regional Crime Center



Data Used in Predictive Policing

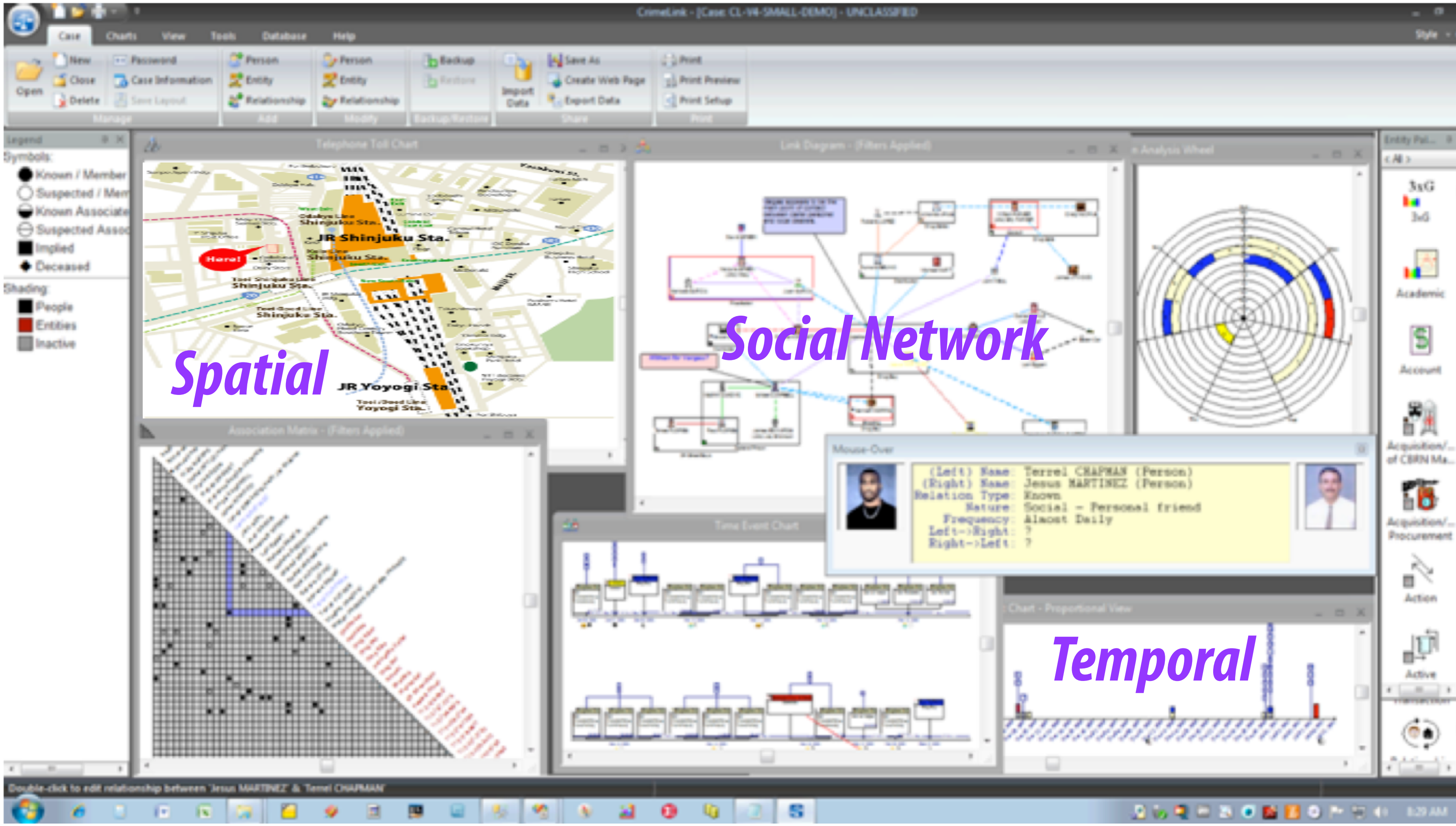
- Three Categories
  - Spatial
  - Temporal
  - Social Network

Spatial Variables	Temporal Variables	Social Network Variables
Indicators of Areas with Potential Victims/Targets • Shopping malls • Property values • Hotels • Area demographics • Population density • Residential instability **Indicators of Escape Routes** • Highways • Bridges • Tunnels • Public transportation • Railways • Dense foliage **Indicators of Criminal Residences** • Bars and liquor stores • Adult retail stores • Fast food restaurants • Bus stops • Public health information • Areas with physical decay	• Payday schedules • Time of day • Weekend vs. weekday • Seasonal weather (e.g., hot versus cold weather) • Weather disasters • Moon phases • Traffic patterns • Sporting and entertainment events	• Kinship • Friendship • Affiliation with an organization • Financial transaction • Offender/victim

The Transition from Data to Knowledge in a Police Agency

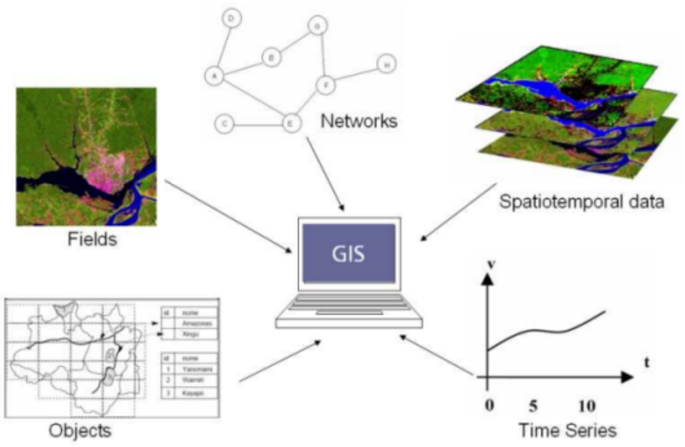
Data		Information		Knowledge		Result
Individual Incident Reports in a records management system	Analysis	Six of the reports are related in a series of robberies	Communication	Robbery series is prime topic of discussion in next detective's meeting	Strategy & Action	Robbery offender is apprehended
Statistics showing number of officers per capita throughout the state		Your police department has 20% fewer officers per capita than average		Chief has this information in mind when preparing his budget proposal		Agency is granted additional officers by city
Crime volume of current year compared to past years; individual records in RMS; jurisdictional information		Auto theft is up 20%, with most of the increase in Police Beat 5 on the midnight shift, probably influenced by new sports arena		Officers internalize this information and consider it when patrolling Beat 5		Auto theft is reduced

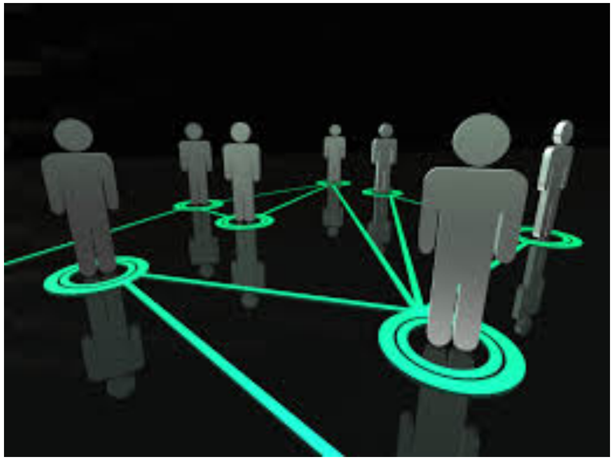
Full Scale Analysis



Predictive Methodologies

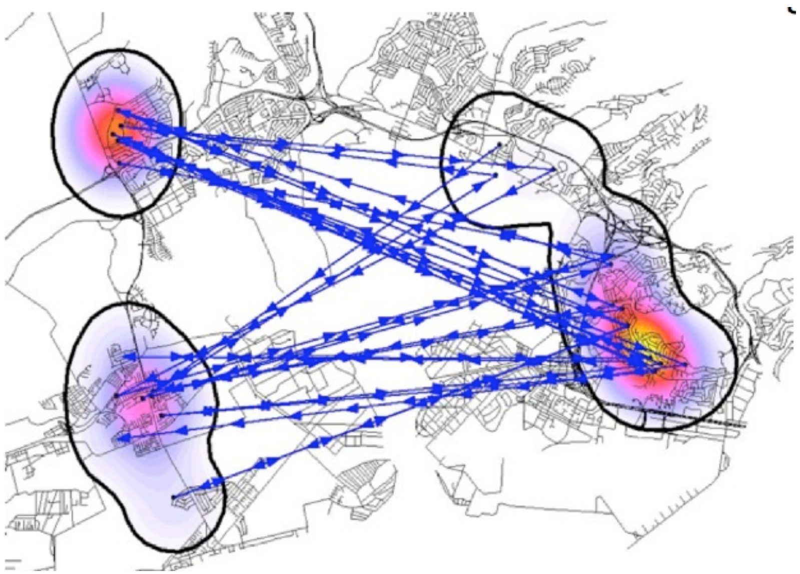
- Analysis of space
- Analysis of time and space
- Analysis of social networks

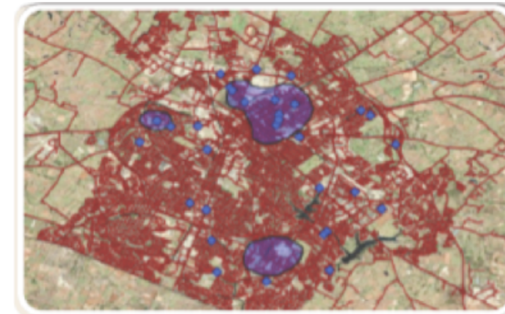




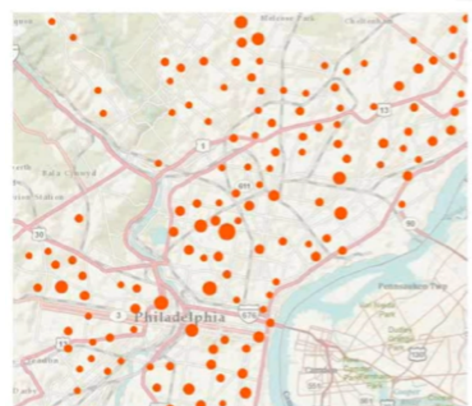
Predictive Methodology One: Analysis of Space

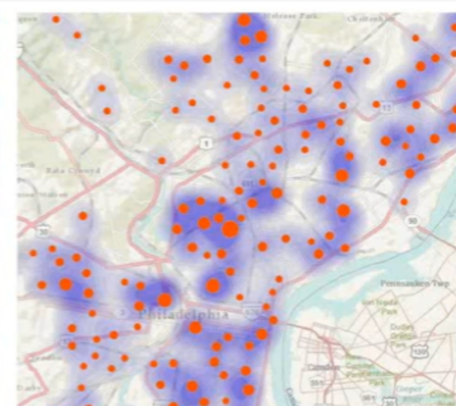
- Point (or offense) locations
- Hierarchical clusters
- Partitioned clusters
- Fuzzy clusters
- Density mapping
- Risk-terrain modeling (RTM) clusters





All crime related gun incidents





Predictive Methodology One: Analysis of Space

- Point (or offense) locations
  - *Theory of repeat victimization, 500x500 feet* [PredPol Software]
- Hierarchical clusters
  - use a *nearest-neighbor technique* [display the clusters: *ellipses, convex hulls]*
- Risk-terrain modeling (RTM) clusters

Gun shootings example

Gun shootings example

Predictive Methodology Two: Analysis of Time and Space

Predictive Methodology Two: Analysis of Time and Space

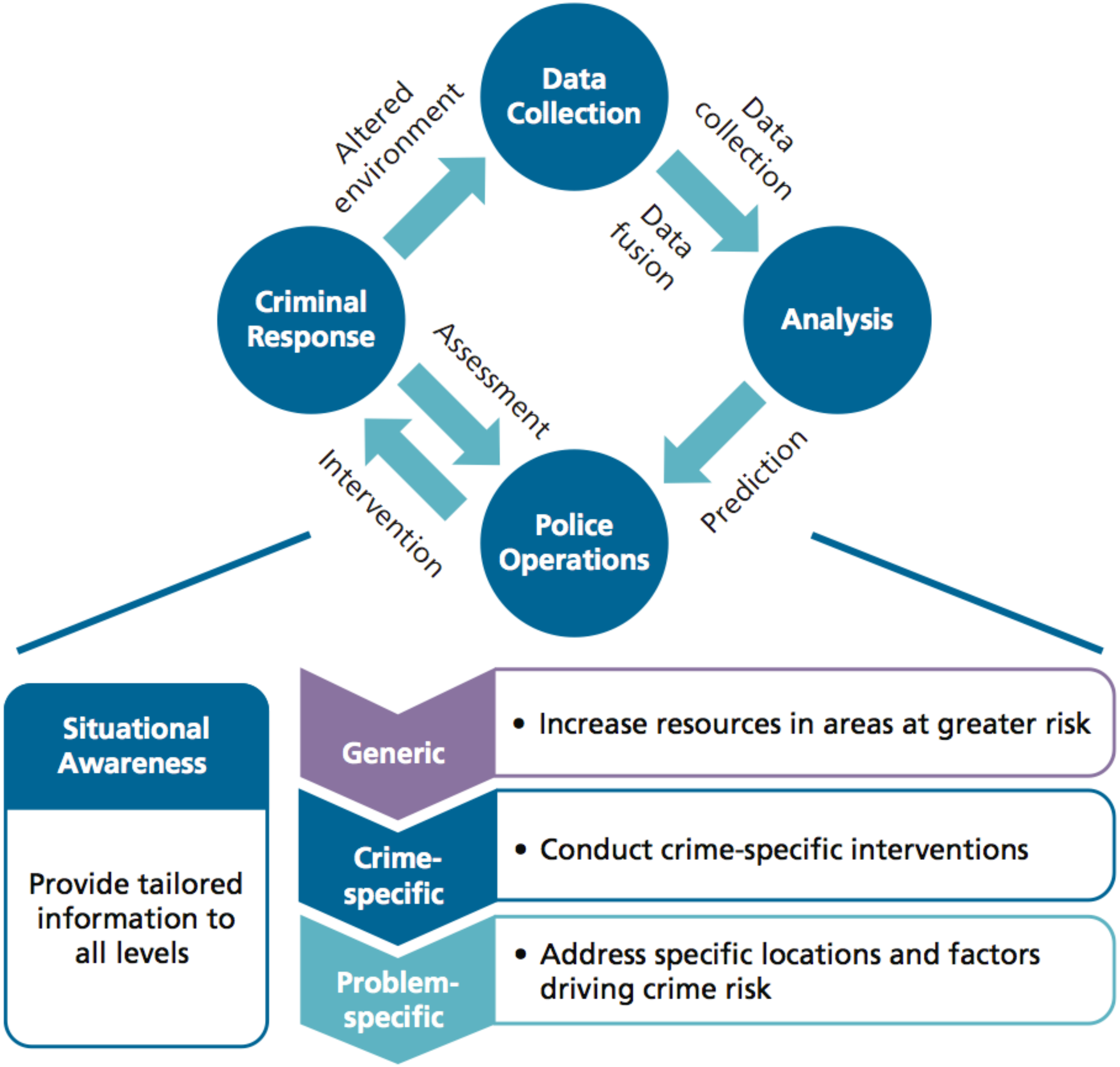
- **CrimStat III : a software program (sociologist + National Institute of Justice)**
  - **spatial-temporal moving average (STMA)**
    - **the *average time* and *location* for a subset of incidents**
  - **correlated walk analysis (CWA) : *temporal* and *spatial relationships* between incidents**
    - **computing the *correlation* between *intervals* [time, distance, direction between two events]**

Predictive Methodology Three: Analysis of Social Networks

- Social Network Analysis (SNA) : *cutting edge* of crime analysis
  - detect *persons* of interest, as opposed to *locations* of interest
  - identify individuals that are central to criminal organizations (eg., gangs and drug networks)

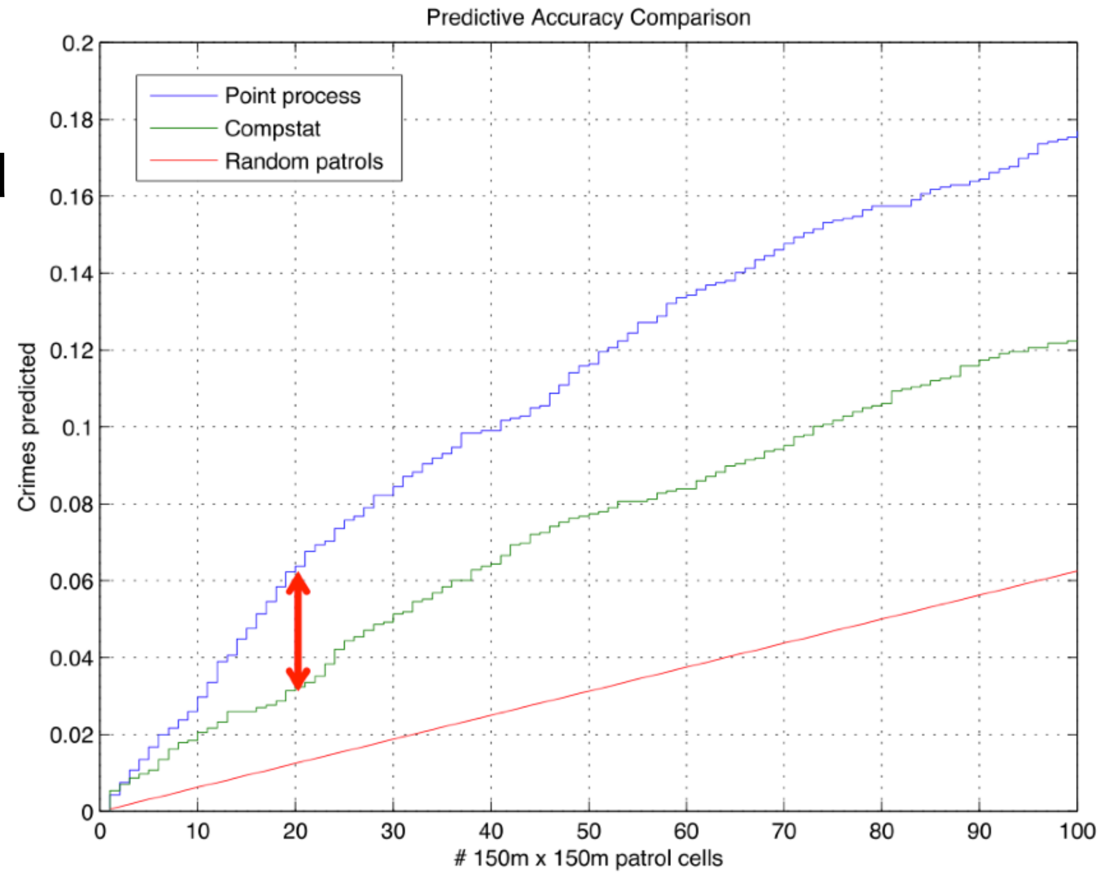
- Building blocks of a social network: *relationships* between two actors
- In crime-fighting applications,
  - SNA is frequently used to identify *central nodes* [*high level of connectivity*]
  - Measures of *centrality*
    - *degree* : *the number of links* possessed by a node
      node's level of *connectedness*
    - *closeness* : *the total distance* from a node to all other nodes in the network
      ease of obtaining information from the network
    - *betweenness* : *the number of instances* a given node appears in the shortest path between other nodes
      relevance to the passage of information within the network

Review: Prediction-Led Policing Business Process [2]



Places on the Frontier of Predictive Policing in the United States

- *Los Angeles, California*
- **Santa Cruz, California**
- **Baltimore Country, Maryland**
- **Richmond, Virginia**
- **Memphis, Tennessee**



Predictive Policing is promising in historical data analysis

Summary

- Predictive Policing *Concept*
- *Mathematical Frameworks* for Predictive Policing
- *Data* Used in Predictive Policing
- *Predictive Methods* for Predictive Policing
- Predictive Policing *Business Process*


рахмат
danke
謝謝
spas
ngiyabonga
Баярлалаа
mersi
kia ora
barka
welalin
tack
teşekkür ederim
mahalo
tapadh leat
спасибо
faafetai lava
vinaka
спасибі
blagodaram
dank je
misaotra
matondo
paldies
grazzi
хвала
kiitos
dankie
nanni
nandri
dhanyavad
thank you
gracias
asante
manana
obrigada
murakoze
tenki
bedankt
bayarlalaa
gràçie
hvala
maururu
köszönöm
akun dankon
ačiū
enkosi
djiere dieuf
tau
mochchakkeram
дякую
mamnun
dziękuję
chnorakaloutioun
gratias ago
gràcies
sulpáy
go raibh maith agat
chokrane
sobodi
dēkuji
sagolun
sukriya
kop khun krap
taiku
obrigado
mèsi
didi madloba
najis tuke
ありがとう
tanemirt
grazie
arigatô
takk
dakujem
trugarez
kam sah hamnida
rahmat
terima kasih
rahmet
diolch
dhanyavadagalu
shukriya
mercé
мерси
তোমাকে ধন্যবাদ
감사합니다
xièxie
ευχαριστώ
merci