

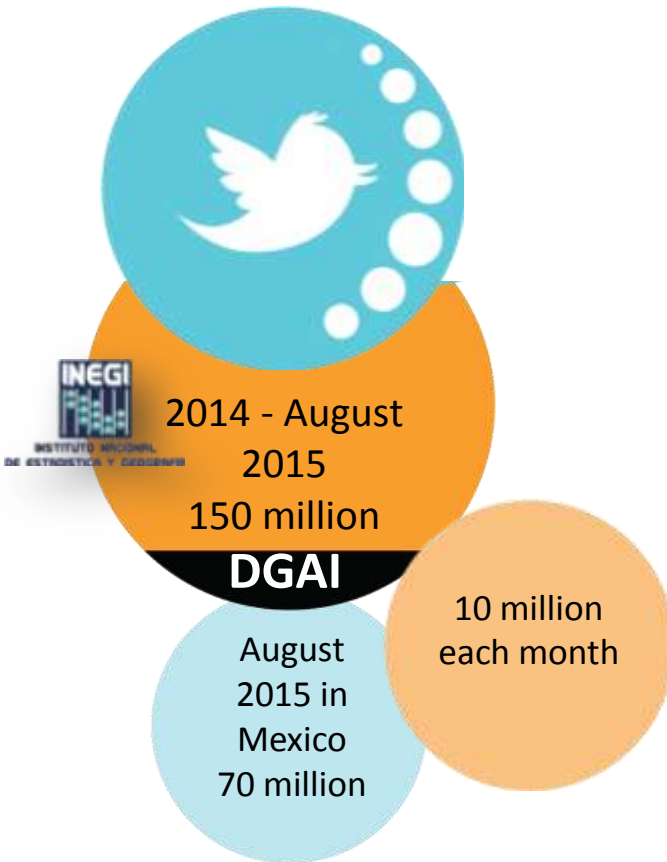


BIG DATA USE FOR LAND MANAGEMENT

Currently is more common to use Big Data in geographic data analysis, for example *Happiness Levels* and *Tourism Routes*.

Twitter is an online social networking service that enables to send and read short 140 character messages called “tweets”, which can be geo-referenced, meaning their metadata includes location information.

INEGI collects Twitter information to get user’s path through Rural Urban System-SUR.





Objective:

Perform an analysis of mobility (source – destination) starting from geographical marks that register users of twitter as part of the study to determine functional dependencies between different SUR communities.

Input

**Tweets on a national scale (DGAI)
Period February 2014-2015**

- Dirección General Adjunta de Informática (DGAI) collects tweets and process them assigning an identifier called FN (Feature Net), which comes from a National Geographic Grid built from to break up the country into small pictures of 5 km²





USING TWEETS ON THE TERRITORY FUNCTIONAL READING



current shape

Rural Urban System:

57	SUR
134	SubSUR
287	CAS
66	CISBA Urban
1,780	CISBA Rural

“Regionalization is constructed from the analysis of circulation networks, flows of people, goods and economic and social information, which is aimed at determining economic structures characterized by convergence of interactions from different orders among the population and sources of employment ¹”.





Methodology

A geographic grid nationwide 5 km² per cell was created

Association tweets to its corresponding box on the geographic grid

Tweets chronological order issued by user

Selecting tweets sequence where a change occurs about the location of this geographic grid squares, to what we define as a path





Methodology

Setting the SUR and Urban locality identifiers to *geographical dependencies* is done with overlapping spatial operation

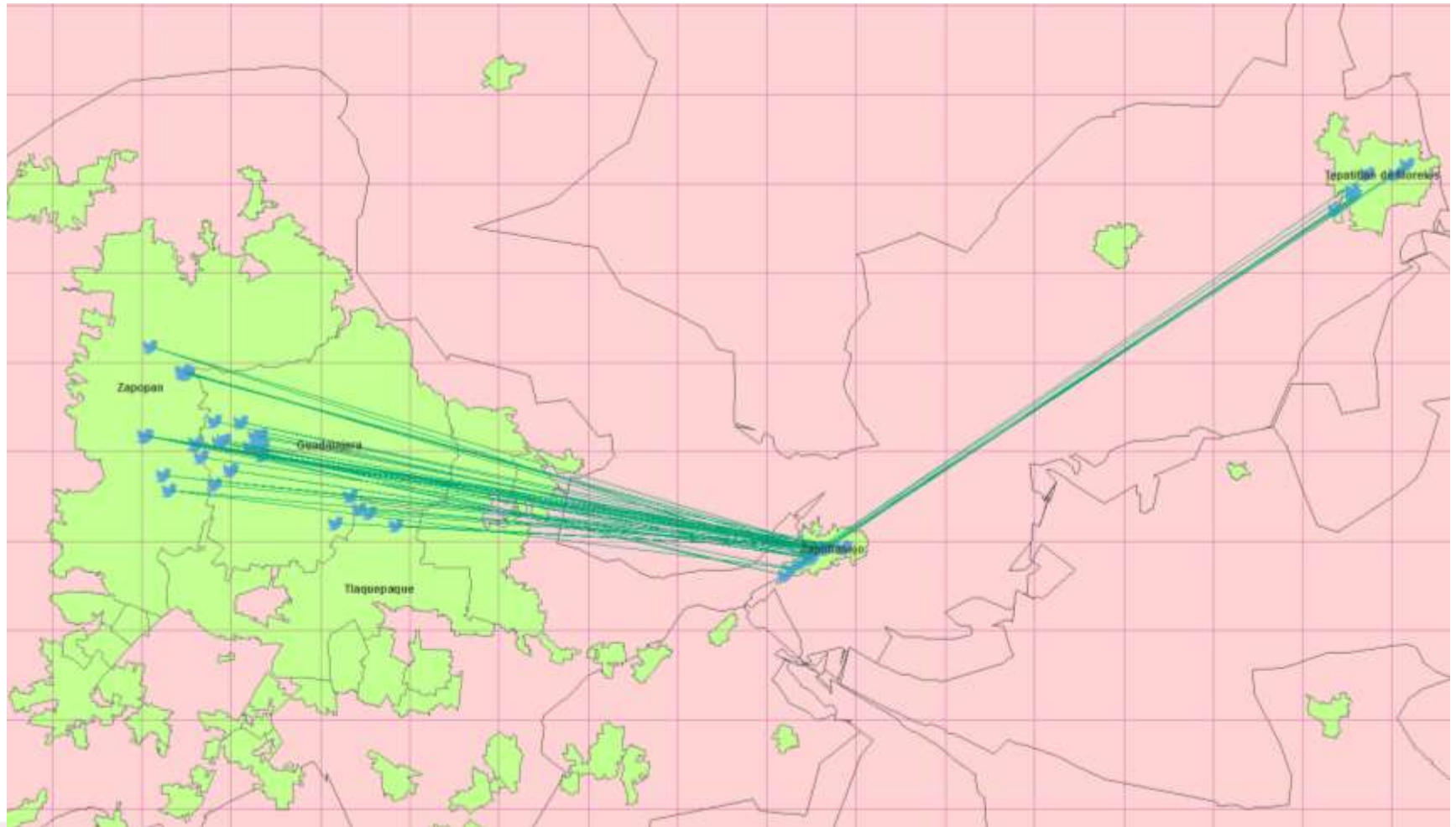
	identificador character varying(50)	FN Inicio numeric(10,0)	FN Destino numeric(10,0)	Localidad Inicio character varying(70)	Localidad Destino character varying(70)	SUR Inicio character varying(30)	SUR Destino character varying(30)
1	01BrissaLombera	60985	46556	Uruapan	Arteaga	MORELIA	
2	01BrissaLombera	64610	60985	Morelia	Uruapan	MORELIA	MORELIA
3	01BrissaLombera	64609	60985	Morelia	Uruapan	MORELIA	MORELIA
4	01BrissaLombera	60985	64609	Uruapan	Morelia	MORELIA	MORELIA





TWEETS IN CHRONOLOGICAL ORDER

- Analysis of tweets per user
- Selection of pairs of tweets that are located in different geographical mesh





Process

Identifying paths and geographical dependencies

Path: Line created by a pair of tweets issued by an user in different location

Dependence: Three or more paths.



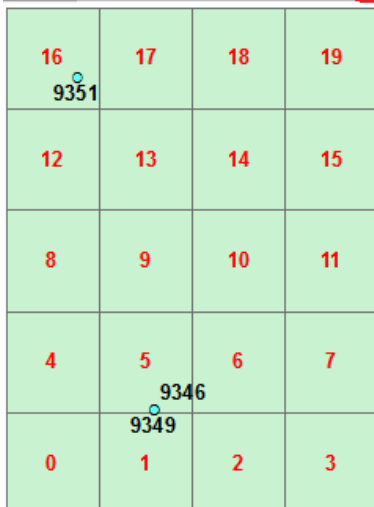


USING TWEETS IN THE STUDY OF LAND FUNCTIONAL DEPENDENCIES

Sequence of tweets per user

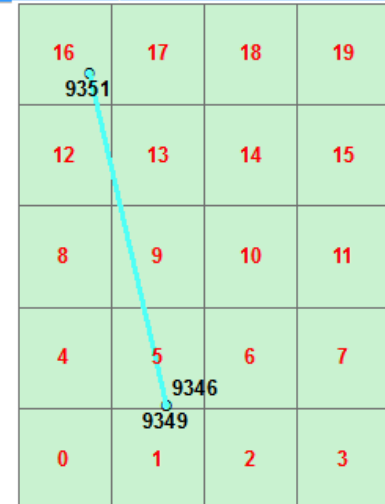
A Trajectory is defined from a pair of tweets issued by the same user and that are located in different frame of 5km².

Data Output												
	Explain	Messages	History									
	gid integer	join_count numeric(10,0)	target_fid numeric(10,0)	id numeric(10,0)	__source_cr character varying(254)	lat numeric	lon numeric	join_cou_1 numeric(10,0)	target_f_1 numeric(10,0)	geom geometry(Point)	fecha1 timestamp without time zone	
1	9349	1		5	1581365028	2014-02-18T14:19:09	20.6174413400	-103.421559300	1	9348	0101000000C7B4	2014-02-18 14:19:09
2	9345	1		5	1581365028	2014-02-24T13:02:29	20.6174413400	-103.421559300	1	9344	0101000000C7B4	2014-02-24 13:02:29
3	9347	1		5	1581365028	2014-02-27T19:20:46	20.6174413400	-103.421559300	1	9346	0101000000C7B4	2014-02-27 19:20:46
4	9350	1		5	1581365028	2014-03-04T14:03:38	20.6174413400	-103.421559300	1	9349	0101000000C7B4	2014-03-04 14:03:38
5	9348	1		5	1581365028	2014-04-03T18:17:33	20.6174413400	-103.421559300	1	9347	0101000000C7B4	2014-04-03 18:17:33
6	9351	1		16	1581365028	2014-04-08T00:56:32	20.7680376000	-103.456957100	1	9350	0101000000330B	2014-04-08 00:56:32
7	9346	1		5	1581365028	2014-04-08T12:03:28	20.6174413400	-103.421559300	1	9345	0101000000C7B4	2014-04-08 12:03:28



Output pane

Data Output	Explain	Messages	History		
	id integer	gids text	fn text	fecha1 timestamp without time zone	fecha2 timestamp without time zone
1	1581365028	9349-9351	5-16	2014-02-18 14:19:09	2014-04-08 00:56:32
2	1581365028	9351-9346	16-5	2014-04-08 00:56:32	2014-04-08 12:03:28





Products

Functional dependencies between SUR and SUBSUR communities

Dependencies among localities in Mexico

Dependencies between SUR and Urban Communities

Dependencies between SUR and Feature Network -FN

Top Ten Urban Communities

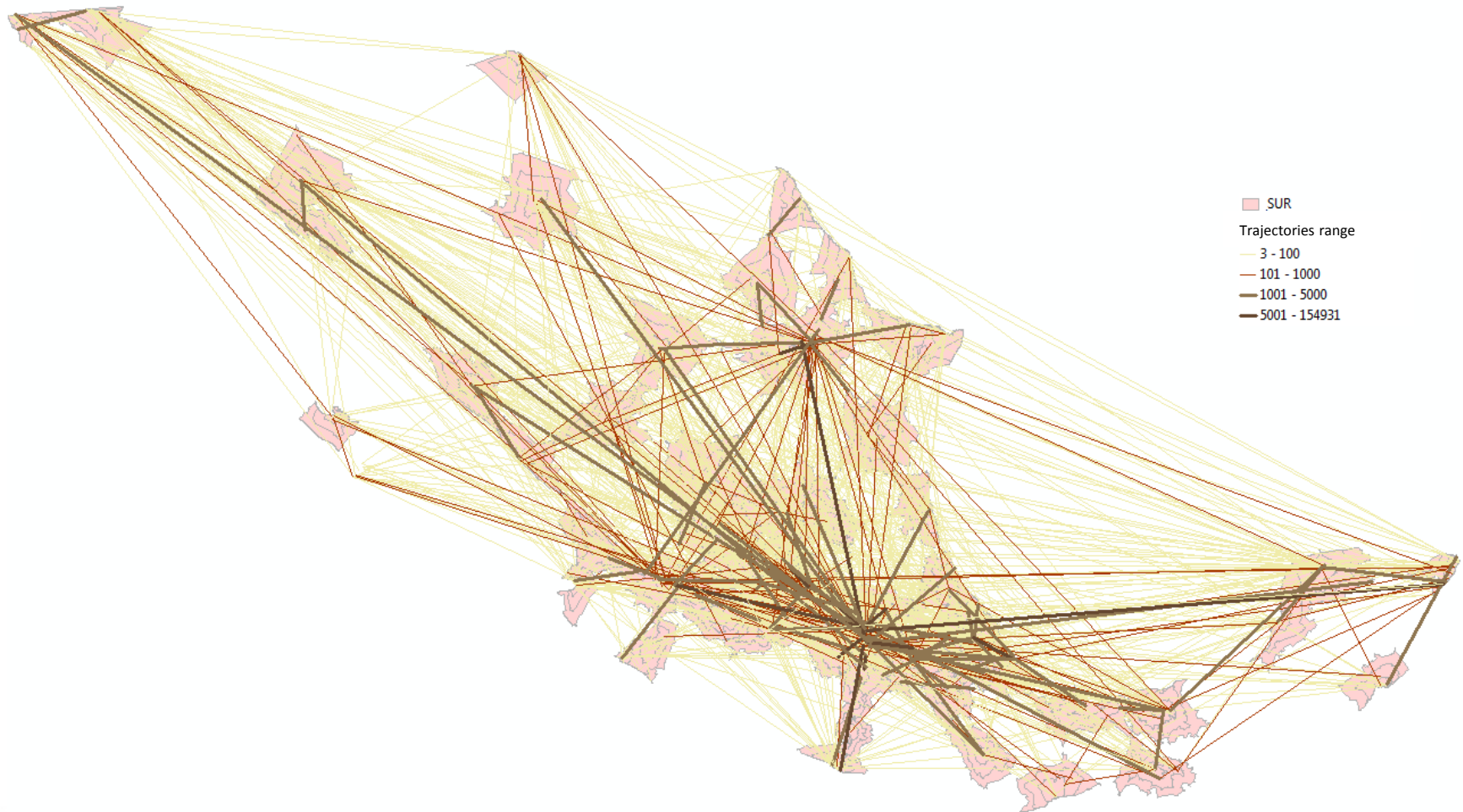
Five communities with greatest amount of dependencies by SUR





DEPENDENCIES AMONG COMMUNITIES OF THE RURAL URBAN SYSTEM (SUR)

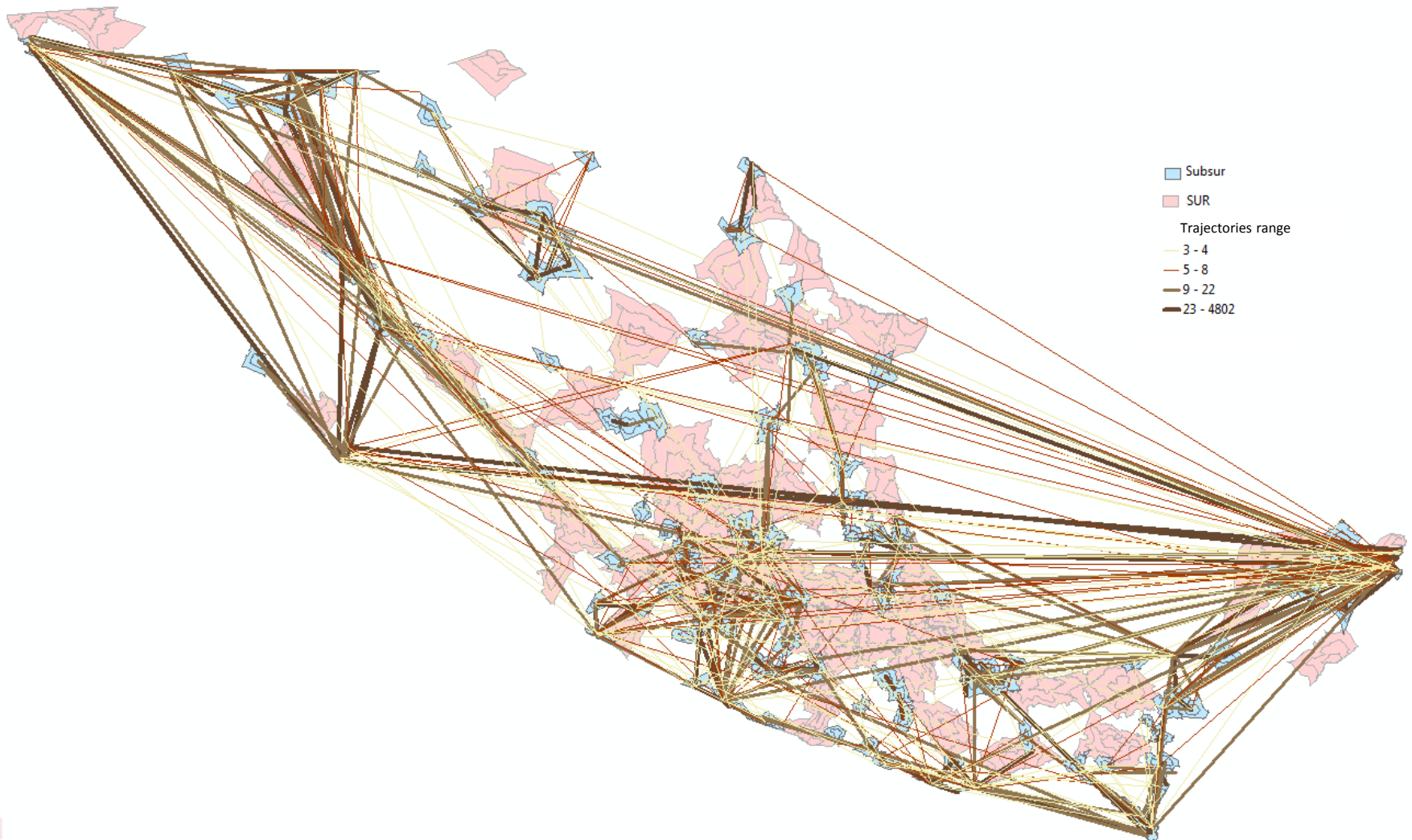
The following chart contains the paths that occur at the national level between each of the polygons in the SUR, which is useful for identifying the dependence between the towns that comprise it.





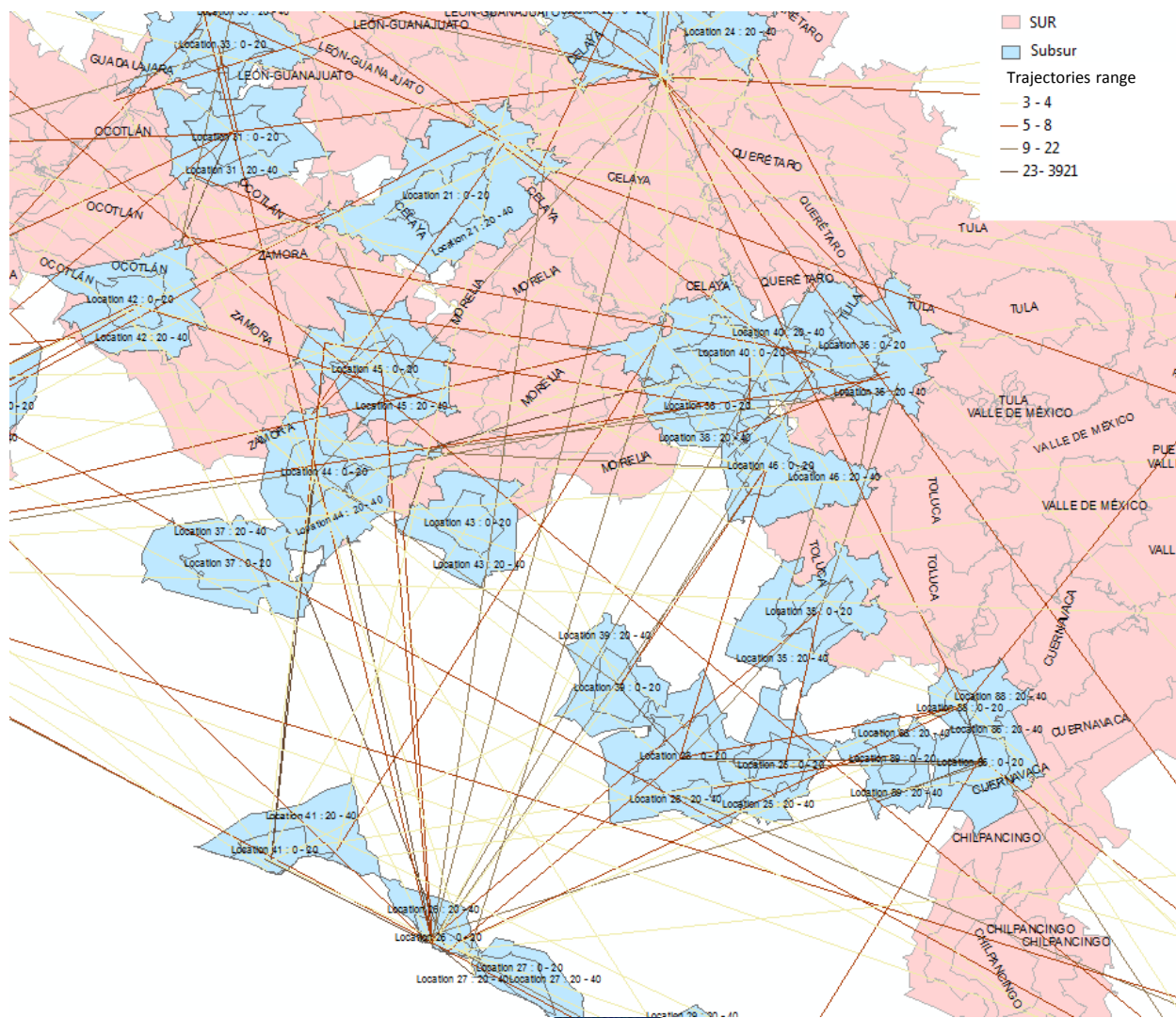
DEPENDENCIES AMONG COMMUNITIES OF THE RURAL URBAN SYSTEM

Territorial Functional dependencies between communities of SUR and SUBSUR is another kind of spatial analysis done from tweets information.





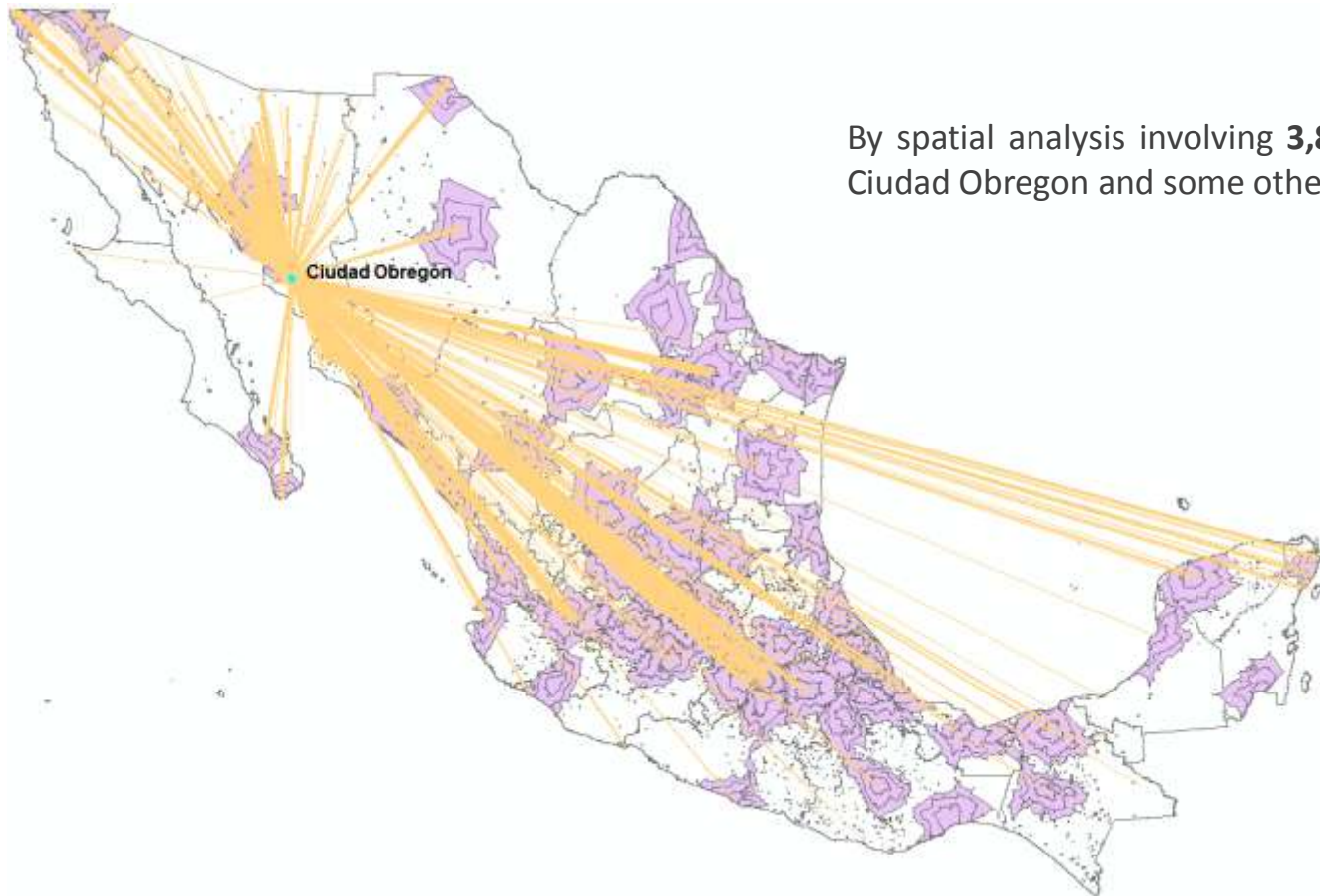
DEPENDENCIES AMONG COMMUNITIES OF THE RURAL URBAN SYSTEM



INSTITUTO NACIONAL
DE ESTADÍSTICA Y GEOGRAFÍA



DEPENDENCIES ANALYSIS AMONG DIFFERENT URBAN LOCATIONS



By spatial analysis involving **3,813** Trajectories between Ciudad Obregon and some other places in Mexico.



INSTITUTO NACIONAL
DE ESTADÍSTICA Y GEOGRAFÍA



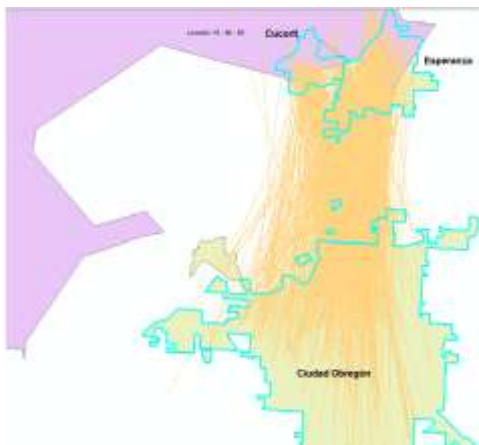
DEPENDENCIES AMONG COMMUNITIES OF THE RURAL URBAN SYSTEM

- Dependencies between SUR and Urban Communities

State	Locality	Key Locality	Number of Trajectories
Sonora	Ciudad Obregón	260180001	3813
	Hermosillo	260300001	3654
	Navojoa	260420001	1523
Campeche	Ciudad del Carmen	040030001	3111
Sinaloa	Los Mochis	250010001	2551
Nuevo León	Linares	190330001	1950
Chiapas	Tuxtla Gutiérrez	071010001	1793
Tabasco	Villa Hermosa	270040001	1694
Yucatán	Mérida	310500001	1677

* Solo se consideraron aquellas localidades urbanas con mas de 1000 Trayectorias

Ciudad Obregón con Cócorit y Esperanza
Trajectories: 2214



Ciudad Obregón con Hermosillo
Trajectories : 884



Ciudad Obregón con Navojoa
Trajectories : 715





DEPENDENCIES AMONG COMMUNITIES OF THE RURAL URBAN SYSTEM AND FEATURE NETWORK

- Dependencies between SUR and Feature Network -FN

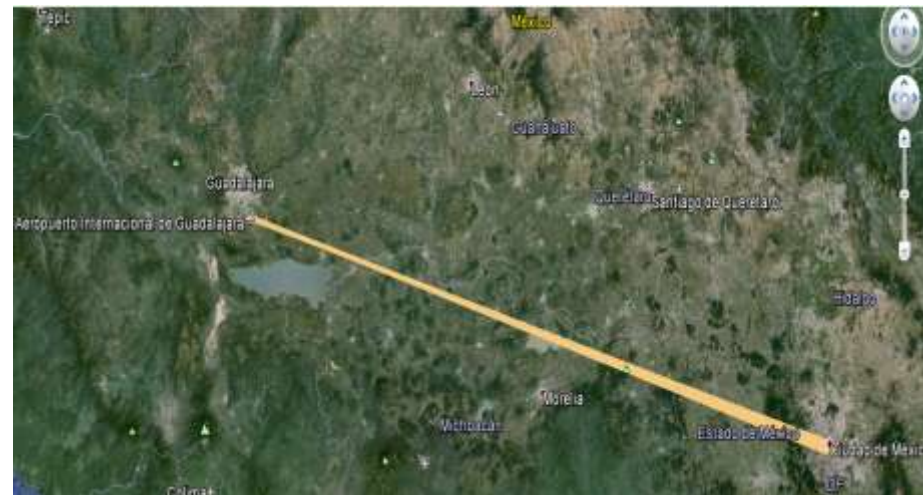
State	Locality	Trajectories	Polygons SUR	Urban Localities
Distrito Federal	Iztacalco	1097	41	Ciudad Apodaca, Nuevo León
		539	41	Guadalajara International Airport



* Only those considered urban areas with more than 500 Trajectories

Urban Localities: Ciudad Apodaca, Nuevo León
Trajectories: 1097

Urban Locality: Aeropuerto Internacional de Guadalajara
Trajectories : 539



INSTITUTO NACIONAL
DE ESTADÍSTICA Y GEOGRAFÍA



TOP TEN URBAN LOCALITY WITH THE LARGEST NUMBER OF TRAJECTORIES

Top Ten Urban Communities (Number of dependencies)

	Trajectories	Locality 1			Locality 2		
		State	Municipality	Locality	State	Municipality	Locality
1	210,291	Distrito Federal	Iztacalco	Iztacalco	Distrito Federal	Coyoacán	Coyoacán
2	148,041					Azcapotzalco	Azcapotzalco
3	146,728	Distrito Federal	Álvaro Obregón	Álvaro Obregón	Distrito Federal	Iztacalco	Iztacalco
4	100,666					Coyoacán	Coyoacán
5	90,615	Nuevo León	Monterrey	Monterrey	Nuevo León	San Pedro Garza García	San Pedro Garza García
6	88,603	Distrito Federal	Álvaro Obregón	Álvaro Obregón	Distrito Federal	Azcapotzalco	Azcapotzalco
7	75,385		Coyoacán	Coyoacán		Azcapotzalco	Azcapotzalco
8	71,997	Nuevo León	Guadalupe	Guadalupe	Nuevo León	San Pedro Garza García	San Pedro Garza García
9	55,448	Distrito Federal	Cuajimalpa de Morelos	Cuajimalpa de Morelos	Distrito Federal	Coyoacán	Coyoacán
10	48,253		Gustavo A. Madero	Gustavo A. Madero		Azcapotzalco	Azcapotzalco



INSTITUTO NACIONAL
DE ESTADÍSTICA Y GEOGRAFÍA



LOCALITIES SYSTEM OF URBAN RURAL WITH THE LARGEST NUMBER OF TRAJECTORIES

Dependencies among different SUR, broken down by locality

Polygon SUR origin	Total Trajectories	Polygon SUR Locality External	Locality External			Trajectories
			Estado	Municipio	Localidad	
TULA	41266	VALLE DE MÉXICO	Distrito Federal	Azcapotzalco	Azcapotzalco	13150
		VALLE DE MÉXICO	México	Atizapán de Zaragoza	Ciudad López Mateos	7575
		VALLE DE MÉXICO	Distrito Federal	Gustavo A. Madero	Gustavo A. Madero	4277
		VALLE DE MÉXICO	Distrito Federal	Iztacalco	Iztacalco	4227
		VALLE DE MÉXICO	Distrito Federal	Álvaro Obregón	Álvaro Obregón	2542
VALLE DE MÉXICO	37624	CUERNAVACA	Morelos	Cuernavaca	Cuernavaca	4779
		TOLUCA	México	Metepec	Metepec	4157
		TOLUCA	México	Toluca	Toluca de Lerdo	3940
		TOLUCA	México	Lerma	Lerma de Villada	3694
		PUEBLA-TLAXCALA	Puebla	Puebla	Heróica Puebla de Zaragoza	3676
PACHUCA	13254	VALLE DE MÉXICO	Distrito Federal	Azcapotzalco	Azcapotzalco	2210
		VALLE DE MÉXICO	Distrito Federal	Iztacalco	Iztacalco	1825
		VALLE DE MÉXICO	Distrito Federal	Coyoacán	Coyoacán	1520
		PUEBLA-TLAXCALA	Puebla	Puebla	Heróica Puebla de Zaragoza	1287
		VALLE DE MÉXICO	Distrito Federal	Álvaro Obregón	Álvaro Obregón	666
LEÓN-GUANAJUATO	9373	CELAYA	Guanajuato	Irapuato	Irapuato	3034
		CELAYA	Guanajuato	Salamanca	Salamanca	1147
		CELAYA	Guanajuato	Celaya	Celaya	909
		GUADALAJARA	Jalisco	Guadalajara	Guadalajara	461
		CELAYA	Guanajuato	San Miguel de Allende	San Miguel de Allende	360
MONTERREY	8662	VALLE DE MÉXICO	Distrito Federal	Iztacalco	Iztacalco	2695
		SALTILLO	Coahuila de Zaragoza	Ramos Arizpe	Ramos Arizpe	1012
		SALTILLO	Coahuila de Zaragoza	Saltillo	Saltillo	776
		VICTORIA	Tamaulipas	Victoria	Ciudad Victoria	453
		GUADALAJARA	Jalisco	Guadalajara	Guadalajara	337
AGUASCALIENTES	1503	GUADALAJARA	Jalisco	Guadalajara	Guadalajara	265
		LEÓN-GUANAJUATO	Guanajuato	León	León de los Aldama	233
		VALLE DE MÉXICO	Distrito Federal	Iztacalco	Iztacalco	167
		LEÓN-GUANAJUATO	Jalisco	San Juan de los Lagos	San Juan de los Lagos	109
		GUADALAJARA	Jalisco	Zapopan	Zapopan	85

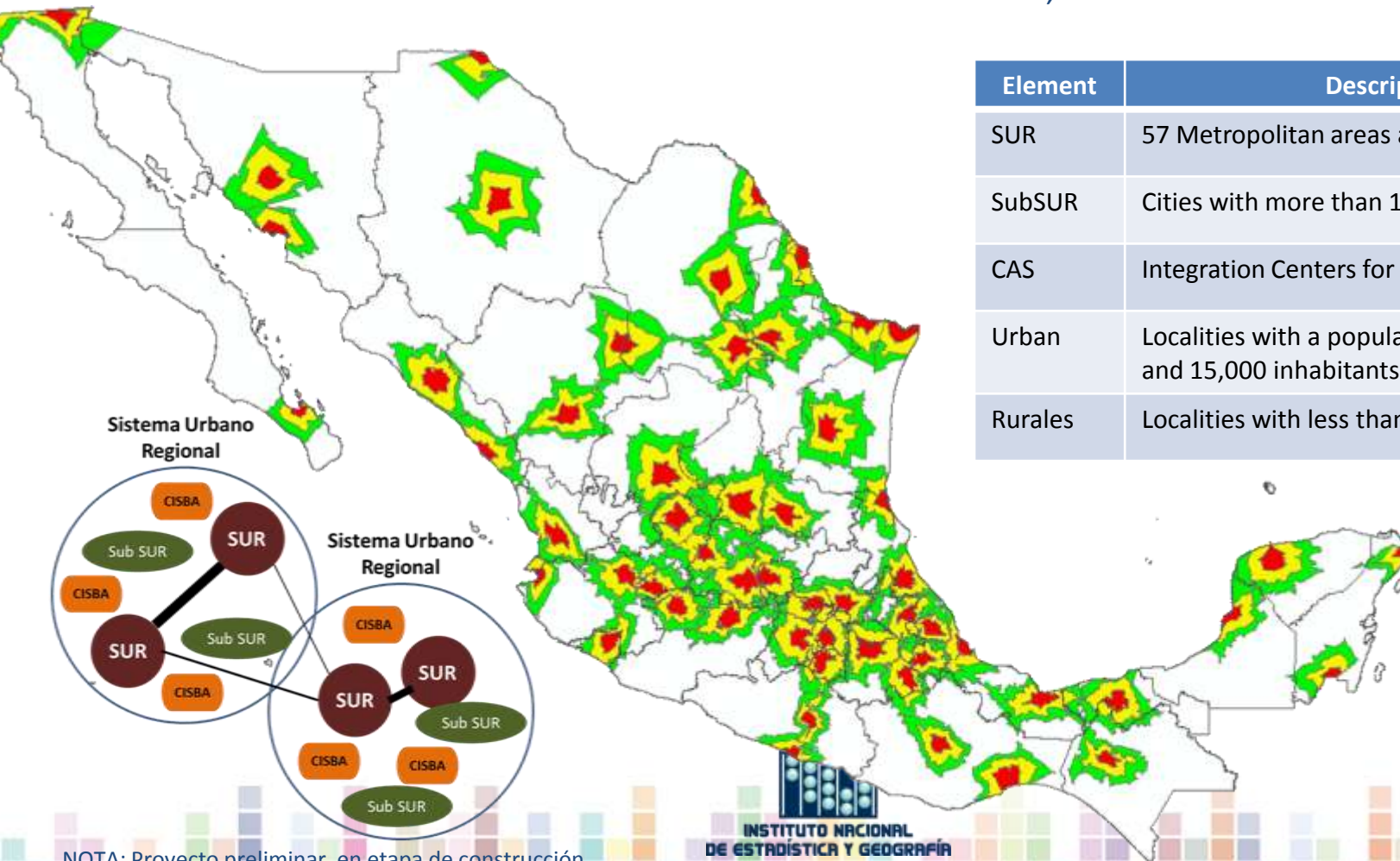


Hoja de cálculo de Microsoft Excel

Urban-Rural System

Territory Functional Reading

The model of rural-urban system, integrating the different settlements of the country, organized by size, geographical distribution, and equipment, services and infrastructure available, is classified as:



Element	Description
SUR	57 Metropolitan areas and capitals.
SubSUR	Cities with more than 15,000 inhabitants
CAS	Integration Centers for Basic Services
Urban	Localities with a population between 2,500 and 15,000 inhabitants
Rurales	Localities with less than 2,500 inhabitants

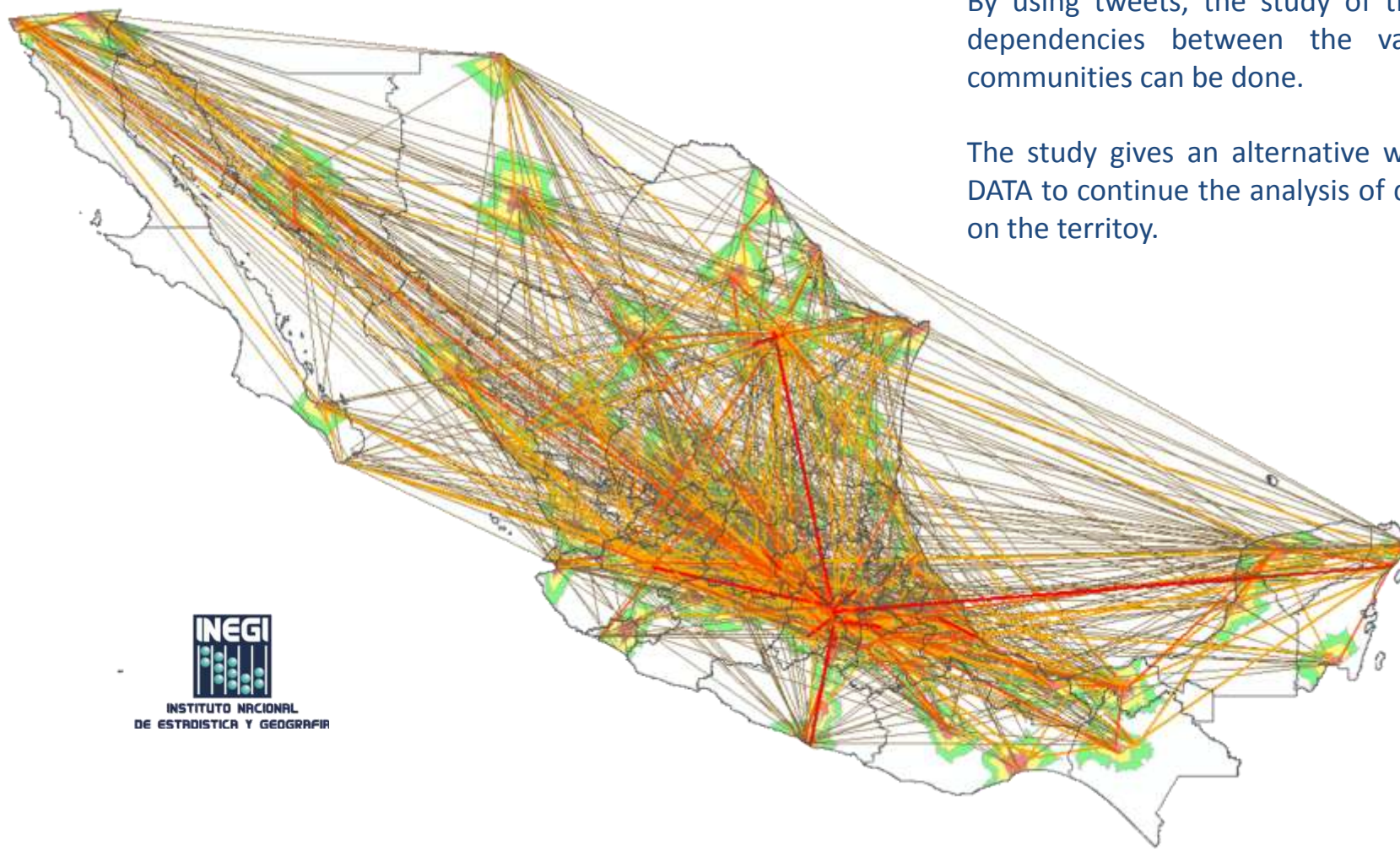
NOTA: Proyecto preliminar, en etapa de construcción.

Using tweets in the analysis of the functional dependencies on the territory

Big Data : Tweets

By using tweets, the study of the functional dependencies between the various urban communities can be done.

The study gives an alternative way using BIG DATA to continue the analysis of dependencies on the territory.



¡Thanks!

Raquel Terán Ramírez

Raquel.teran@inegi.org.mx



INSTITUTO NACIONAL
DE ESTADÍSTICA Y GEOGRAFÍA