City harnesses big data to study crime

Advanced analytics tool spots trends

ELISE STOLTE
Edmonton Journal

Social workers and neighbourhood police officers hope a new data tool will clarify the characteristics in each neighbourhood that are inviting property or violent crime.

The new interactive dashboard, based on 233 potential factors, uses advanced computer analytics to drill down to a level roughly the size of a city block.

“Now the downtown Neighbourhood Empowerment Team, a crime-prevention unit composed of police officers and social and youth workers, will test the data in a year-long pilot project starting in January,” said Kris Andreychuk, head of the non-profit neighbourhood safety organization.

“I think it fundamentally shifts the way we look at crime. It’s not about bad people,” said Andreychuk, who supervises the crime prevention team. “It’s a conversation changer.”

Andreychuk and Stephanie Contra, the city’s senior information architect, started eight months ago with a literature review that found 18 factors thought to cause crime, including “rowdy teenagers.” They pulled in experts from the University of Alberta, MacEwan University and the Edmonton Police to expand the list with factors such as legging areas, alleys, pawnshops and body-rub centres.

Using that as a guide, they found 233 geographically based data sets among them, city light poles, noise complaints, school sites, picnic sites, front yards in bloom awards, retirees, vacant buildings — from sources that included Statistics Canada, police, 311, the Yellowpages and other city departments.

To analyze the data, they used a base layer of crime statistics, and then used a computer program to systematically attempt to match the crime statistics with all of the other data sets, on their own and in combination, to see which ones could be statistically correlated to high or low crime.

That generated 92 rule sets that predict the presence of crime. But it’s still pretty high-level stuff. It turned out, nothing is significant enough to be a predictor of crime on its own.

But, for example, rule set 43 says that 100 per cent of the time you can find a high incidence of property crime in those areas where you find:

• a high number of recovered stolen vehicles,
• plus a high number of noise complaints,
• plus a high concentration of youth services, and
• combined with a low concentration of picnic sites/fee pits.

“I would have never thought that picnic sites have skin in the game,” said Andreychuk.

There’s a lot of work to be done before anyone fully understands what each of these rule sets means, whether stolen vehicles are important on their own or if they simply signify an area where neighbours don’t know each other.

But if it’s a neat improvement over acting only on the basis of where crimes have happened in the past, Andreychuk said, it will be a starting point for a long conversation with members of the community where the crime is occurring.

This type of analytics is at the forefront of new policing models, said Capt. David Veitch, head of the co-ordinated policy division, which includes several units considering new risk-based approaches.

“Math and crime are closely associated today,” Veitch said. “In Edmonton, police have seen good results with a heat map approach, mapping crime in almost real time and increasing their presence where problem spots appear.

But in the United States, police are starting to incorporate predictive police models. Rutgers University in New Jersey developed a system called “risk terrain